

After Windscale, what?

THE report by Mr Justice Parker into British Nuclear Fuels Ltd's application to extend reprocessing facilities at Windscale breaks much new ground. Not only was the inquiry on an exceptionally broad scale, bringing into an ostensibly local affair all the questions of national interest, but now the government has had to find a way of opening the report up for public debate without seeming to contravene the principle that Mr Peter Shore, Secretary of State for the Environment, should, as is normal in planning inquiries, make up his mind on the basis of his inspector's report alone, free from any further hectoring, adducing of 'new evidence' and so on. This particular circle has been squared by the subterfuge described on page 117; Mr Shore has said that it is a cogent and persuasive report but has gone against its line of argument in rejecting the application, simply to allow parliamentary debate.

But, as pointed out on page 115 this is not the end of a planning road so much as the beginning of one. Large-scale investment in energy facilities is just round the corner for all manner of technologies—alternative as well as traditional—and in each case the questions to be raised stretch to the limit the forms of inquiry and the conventions applied to them at present. And not only must there be doubts about the shape of future inquiries, there must also be serious questions about the extent of public participation. The British public—even those who would regard themselves as 'intelligent laymen'—is not particularly well informed on matters of energy and energy policy, in part because of the compartmentalisation of British society, in part because there has been a long-standing trust in the quality of the expert. As yet we not only do not know what the intelligent layman thinks about nuclear issues, we do not know whether the intelligent layman even wants to express a view on these matters.

The next nuclear decision to be taken will come within a year when the Atomic Energy Authority puts up a proposal for the building of CFR1, a demonstration commercial fast breeder reactor, which is possibly planned for the Windscale site. The planning inquiry format, such as was followed at Windscale, is ideal, no

doubt, for purely local matters but is not ideal when matters of national interest are also being debated. Fortunately there already exists on the statute books the concept of the Planning Inquiry Commission in which, say, five commissioners, not necessarily with legal training, hear evidence (including the cross-examination of witnesses) and can initiate their own research. It is envisaged that Planning Inquiry Commissions would consist of two stages, the first of which looked at the broad issues of whether the need existed, the second of which looked at the local impact. The Roskill Commission on the Third London Airport was the closest in style so far to a Planning Inquiry Commission, but one is almost certain to be called over the exploitation of the Vale of Belvoir coalfield.

It can be argued, however, that even a Planning Inquiry Commission does not do the job adequately, as it is dissolved on completing its work (the Minister is still, of course, at liberty to reject its recommendations). What is needed is some sort of standing commission that does not have to be re-educated each time a new decision has to be made—and yet it must not become a group whose views become hardened and utterly predictable. (It should be added that the recently announced Commission on Energy and the Environment is specifically excluded from studying individual planning applications so is not an appropriate body.)

But what about public participation, particularly in the question of CFR1? Commissions may be fine in many respects but they do consist largely of experts talking to each other. There is a strong case for the Commons Select Committee on Science and Technology picking up this issue in order to try and stimulate more parliamentary interest in energy matters than has so far been apparent. Certainly it would prove the most difficult venture the committee had yet undertaken and certainly its deliberations could hardly be a replacement for the standard planning procedures. And yet, if it proceeded with some dispatch it might well provide much of the relevant background, and if hearings were conducted imaginatively it could well play a major role in raising the interest of public and politicians. □



Science and technology for development: no escaping the politico-economic context

Dr Narendra Singh, President of the Indian CSIR Scientific Workers' Association, argues that fundamental political problems must be solved before science and technology can make a contribution to development.

MANY scientists will first have heard of the United Nations Conference on Science and Technology for Development (UNCSTD), through Moravcsik's recent *Nature* article (24 November, page 288). "Will the voice of science in the Third World be heard?" I have basic differences with the elitist features in Moravcsik's article, but I agree with him that the scientists in Third World countries must seriously participate in the formulation of national papers and not leave the task to a small group of bureaucrats and technocrats. In the absence of broad participation, the basic issues may be missed altogether, with serious implications on the national scene and for cooperation among Third World countries.

Perhaps I should first discuss this elitism I have mentioned above. Moravcsik urges Third World scientists to make prompt and forceful inputs in three main areas for incorporation in the national papers, namely the special needs of foreign-educated indigenous scientists, the expansion of communication facilities and evolution of institutions dedicated to long term, not just immediate research, and the practice of science for the purpose of establishing more equality not just in GNP per capita but also in countries' self-images of their standing in the world of science. These points seem quite innocuous and may readily appeal to all concerned, but they do ignore the historical antecedents of industrially developed countries and the way in which science developed therein; they also ignore the present politico-economic context of the Third World countries. Such proposals make no contribution to socially useful science and technology for development in the Third World countries and would ultimately even fail to achieve their espoused purpose.

Take, for example, China. Without paternalistic advice and elitist and esoteric diversion, China has fruitfully applied science and technology to solve the immediate problems of its people and has also made important contributions on the frontiers of science. Fundamental science is the foundation of both immediate and long-term research, and is not incompatible with a balanced programme of science and technology for development. But first the basic political and economic objectives must be clear, and be actively pursued.

National papers of Third World countries should thus enunciate the basic objectives of science and technology for development and then identify areas of action for achieving those objectives. No doubt competent analyses could lead to the identification of obstacles such as the practices of corporations and monopolies, and to proposals that would decrease dependence. But for real progress towards self-reliance, much more than this is needed. It is vital to determine how to change the politico-economic context of operation of those forces which give rise to obstacles to progress. Unless that is done, we continue to grope in a maze, proliferating technology and management-oriented solutions to problems of science and technology for development, but to no avail, since the problems are essentially politico-economic.

Take the case of science and technology for integrated rural development, a subject of great concern at present. The large-scale agro-industries, concentrating on single crops, are proliferating. They depend largely on technology imported from the developed world, and sell principally to local affluent markets and for export.

Such an analysis would frequently be followed by calls for appropriate technologies to decrease dependence, and experts would recommend a separate department of agro-industrial science and technology, the expansion of relevant activity in existing organisations, an integrated technical information centre and the strengthening of relevant technical facilities for appropriate agro-industries. Apparently sound suggestions. But what about those factors which have been promoting technological dependence? Unless there is action to check the operation of those factors, no practical progress is possible. Thus a programme for integral rural development must combine suggestions of measures for eliminating obstructive forces and creating a favourable environment for developments in positive directions together with concrete proposals on appropriate technologies.

The basic objectives of science and technology for development in the Third World countries are economic independence, self-sufficiency in basic needs of the common people and self-reliance in development. The forthcoming UN conference professes the aim of national self-reliance to the extent called for in the new international economic order. The qualifying phrase is unnecessary and is probably made to adjust Third World countries to the modern neo-colonial international division of labour.

No real national self-reliance can be achieved in the politico-economic context of the modern world unless it is linked with the aims of economic independence and self-sufficiency in essentials. 'Economic independence' does not mean economic isolation, but a national capability for economic relations with others on equal terms and from a position of economic strength, instead of the present recipient-donor relation with the industrially developed countries.

'Self-sufficiency' in essentials, in food and others, is basic to economic independence of the Third World countries and to the welfare of the common people there, but this does not necessarily mean a self-sufficiency in each individual item all the time, only that in compelling circumstances there must be enough surplus of some items to trade for other items in need. Likewise, 'self-reliance' in development does not mean self-imposed isolation, but only national capabilities to be able to carry on development pursuits, including those in science and technology, even in the most adverse circumstances of enforced isolation.

The forthcoming UN conference must be oriented towards achieving real national self-reliance, which in turn is not possible without progress towards economic independence and self-sufficiency in basic essentials. The input in national papers must take the basic issues into account and develop them with concrete proposals, instead of mere concern with details of technology and organisation. And these basic issues must be raised for serious discussion at the conference, to encourage cooperation within the Third World and with industrialised countries. In my opinion, these issues are of real significance for Third World participants and must be of concern also to others who are sincere in their professions of an interest in science and development. □

UNCSTD prepares for action

A preparatory committee met in Geneva recently to decide on the agenda for UNCSTD. Peter Collins reports

SCHEDULED for Vienna in August next year, the United Nations Conference on Science and Technology for Development, UNCSTD, could be a milestone towards bridging the ideological gap between the countries of the northern and southern hemispheres. These are the two camps in which the Member nations of the UN seem perpetually divided: the 'Group of 77' (now in fact 103) developing countries and the western and eastern industrialised countries.

Governments are the focus of the preparations for UNCSTD. They have been involved largely through the organisation of seminars at every level from the national or sub-regional to full scale regional meetings. These have been organised or supported not only by the UN itself or such bodies as UNCTAD, but in many cases also by UNESCO. What this sometimes rather frenzied preparatory activity will lead to is not clear, but the Secretary-General of the Conference, Mr Joao da Costa, has already said that he considers the preparatory period as part of the whole Conference procedure. This should mean that when delegates finally get to Vienna, they may have a reasonable idea of what they want to say, or expect to hear—which has not always been the case with such mammoth meetings in the past.

Meanwhile, as a result of the recent meeting, the agenda for Vienna are now firmly established, with five sub-

stantive items for discussion: the part of science and technology in development; suitable institutional arrangements and possible new forms of international cooperation in the application of science and technology; the ways and means of using the UN system in this field; science and technology for the future; and a programme of action based on the conclusions of the Conference.

The original intention was to limit discussion under these headings to not more than five 'subject areas'. However, so much confusion seems to have arisen in this respect, and the proposals from many developing countries have been so diffuse, that no such neat arrangement has been possible. At their first preparatory meetings, the five regional Economic Commissions agreed on priority subjects and a first task at Geneva was to accommodate their ideas for the Conference.

The result is a group of subject areas so broad as to make discussion in depth of any one subject almost impossible: food and agriculture; natural resources including energy; health, human settlements and environment; transport and communications; and industrialisation including the production of capital goods. Each of the first three is further divided into four sub-topics, in such a way that there may be no aspect of development, discussion of which may not



This logo designed by Rashid-ud Din of Pakistan has been adopted as the official symbol of the Conference. It depicts the globe in the form of a retort held by calipers and surrounded by the olive leaves of the United Nations emblem

consume time at Vienna. The result is a list that differs very little, in effect, from that covered by the UNCSAT Conference of 1963, except that there is no special mention of education and training, often considered the most essential prerequisite for the useful application of science and technology anywhere.

Action plan

The item on which the preparatory meeting most nearly broke down was one on which nothing definite can be said until the Conference itself, namely the 'action plan' which, ever since Stockholm in 1972, has become a *sine qua non* for any major international gathering of this type. After a long and apparently quite acrimonious discussion, the Group of 77 came up with a draft resolution, not so much about what they felt the action plan should contain, as about the obstacles to development which it

Sorry, for copyright reasons some images on this page may not be available online

should be concerned to solve. In this text they allowed that many of these obstacles, at the national level in particular, are inherent in the undeveloped condition, and are no longer, as has often been claimed in the past, just the result of colonialism or the behaviour of the multinational corporations. Even so, the text eventually submitted was unacceptable to the western group. However, a revised version put forward by the Conference Secretary General himself was more satisfactory; almost everyone was able to agree to it and all was well, at least for the time being.

National papers

From a practical point of view, the way is now clear for the real preparations to start. What this entails is indicated by the scheme of documentation presented to the recent meeting. Basically, this involves the preparation of national papers, which have been defined as "national analyses of relevant socioeconomic problems which may be solved with the help of science and technology". Together with equivalent papers at regional level, they will form the backbone for the Conference documentation. This falls into five categories. The first comprises the so-called 'official' documents, prepared or commissioned by the Conference secretariat, and covering the five major items of the agenda detailed above.

The most interesting of these to the pure scientist will be the report on Science and Technology for the Future. The recent meeting was shown a paper on this subject prepared by a group of advisers invited by da Costa, but it did not make a particularly good impression and it is not clear who will produce the final version. This is seen as a book of 200 to 300 pages, equivalent to the Barbara Ward and Rene Dubos volume produced for the Stockholm Environment Conference, and available for the general public. At present, da Costa admitted, there are no funds to cover its publication within his budget for the Conference.

In the second category come the national and regional reports, and three important papers, specifically commissioned at the request of the Preparatory Committee. They deal with: the present role of the United Nations system in science and technology; the human rights aspects of science and technology; and an evaluation of the results achieved so far from previous UN conferences in the scientific and technological field.

Then come three groups of background documents: first, those from the UN Specialised Agencies and other members of the UN system; then, re-

ports from inter-governmental organisations (such as OECD and EEC); and lastly, the submissions of the principal non-governmental organisations—The International Council of Scientific Unions (ICSU), Pugwash, and the like. These last could well be particularly valuable from the point of view of the scientific community, since it is in them that scientists will be able to put their points of view, free from the restraints imposed by government policy. Despite assurances that there would not be a vast mountain of documentation for delegates to consume, even the basic 'official' papers will amount to some 700 pages, and the grand total, at a rough but conservative estimate, will be around 5,000!

Potentially, the group of reports from the components of the UN system should be of the greatest interest to the developing countries. After all, it is they who, together with the bilateral agencies, have been applying science and technology to the development of the Third World for the past 25 years, and the sum total of their experience has never yet been made freely available. It is here that the political divisions within the system are most regrettable.

The Specialised Agencies, other than UNESCO, appear to be taking a disinterested attitude towards the Conference, whereas UNCTAD and UNIDO have from the start been more closely and deliberately involved—admittedly, at the express request of the Secretary General, who has turned to them for assistance. The result has been that many governments of the developing countries have come to regard the Conference as primarily, if not exclusively, concerned with the transfer of technology.

French opposition to science

As to the Conference itself, there seems to be a large measure of agreement with da Costa's insistence that it is essentially a political affair. What many people may not realise is that it is just for this reason that the voice of the scientific community must be heard. Too often, major political decisions have been taken at large scale conferences sponsored by the UN, only to prove unworkable when it has come down to the hard facts of putting them into practice. Yet there is still considerable opposition to the idea that science as such should have a voice at Vienna.

The French, for example, are among those who would remove item 4 (Science and Technology for the Future) from the agenda completely, nor are they in favour of the non-governmental organisations being in-

involved in the Conference. No one, of course, expects a series of decisions that can bring about immediate changes in the developing countries. The Conference, as the United States representative put it, is the commencement, not the culmination, of UN's new approach to the problems it will discuss, and any resolutions taken will be framed in that light.

Scientist joins secretariat

While the governments interested in the Conference will be preparing their national papers, deciding what it is they really want from Vienna in 1979, the secretariat itself has a long and tough task ahead. One bonus is that Guy Gresford, at present senior scientific adviser to the Department of Foreign Affairs in Canberra, will shortly join the secretariat as deputy to da Costa.

Gresford's appointment not only means that his long experience and diplomatic skills will be available to the hard-pressed secretariat in New York (and thereby raise their apparently somewhat low morale), but also should ensure that the point of view of science and technology will not be lost in the welter of political manoeuvring that seems inseparable from all such meetings.

Moreover, it seems likely that Gresford, as a former member of the UN top scientists' advisory body, ACAST, and as a person of many years experience inside the UN itself, will be in a position to see that something is done about the greatest gap in the Conference preparations to date. This is the lack of public information about every aspect of the Conference, not only to mass media but also, as several delegates at Geneva remarked, to governments preparing to take part in the meeting.

Even the Secretary General observed that he had been poorly served in this respect, and the Specialised Agencies seemed to be doing a better job than the people in New York.

One activity that may eventually make some difference in this respect is the colloquium proposed by ICSU, to be held if possible between the third session of the preparatory commission, in September this year, and the fourth and final session, in February 1979. This should not only make it possible to clarify and formulate the views of the scientific community well in advance of the actual meeting, but may also provide an opportunity for the kind of publicity that is going to be sorely needed if what happens in Vienna later that year is not going to be dismissed, by the media and the public they serve, as "just another vast United Nations jamboree". □

Energy: how to decide

Planning inquiries, such as the Windscale Inquiry in which the UK debated whether to allow expansion of a nuclear reprocessing plant, are not the only means for debating the UK's technological future. Here **Professor David Pearce, Ms Lynne Edwards, and Mr Geoff Bueret** discuss the options

NUCLEAR power is the technology which could probably meet the 'energy gap' between projected demand and supply most quickly; but it raises issues that a concerned society must debate: the threat to civil liberties, the proliferation of nuclear weapons, the risk of accident and the extraordinary problem of storing radioactive waste for longer than any single civilisation has yet survived.

Science, we are told, must be for the people; so decisions about new technologies must be accountable to the public. But the public has limited technical understanding; and the institutions which exist to resolve public issues are the products of an age when few were concerned with technical problems. In the UK, the standard response has been to rely on extra-parliamentary institutions such as Royal Commissions, major Inquiry Commissions, the institution of the Ombudsman, administrative tribunals, and public inquiries.

But is this an efficient context for deciding on an energy future?

The Windscale Inquiry

Typically, local planning inquiries in the UK consider local issues, although national concerns may be involved: any motorway, for example, is part of a national network so that the local decision affects the network plan. But by and large, the nation's needs never dominated local inquiries—until the inquiry at Whitehaven between June and November 1977 into the proposal to reprocess nuclear waste fuels at Windscale. The Secretary of State for the Environment issued national terms of reference for that inquiry. They spoke of "the safety of the public" and vaguely of "other aspects of the national interest" as well as local issues of employment and amenity.

It is arguable that the national issues raised at the Windscale Inquiry should have been debated elsewhere. Clearly, they should be debated in Parliament; yet by placing the debate in the context of a local inquiry, the Secretary of State appeared inadvertently to obstruct this. Planning inquiry law would be contravened if in the course of a parliamentary debate 'fresh evidence'

was introduced which affected the Minister's own internal assessment, so rendering a proper parliamentary discussion impossible. But the minister has now found an awkward way round this obstacle. There will be a debate.

By including national issues in the terms of reference of the Windscale Inquiry a precedent has been set. Strictly speaking, 'terms of reference' do not exist for local inquiries. In future, therefore, unless the law is changed, any local inquiry could be opened up to debate national issues. This could mean an inquiry every time a nuclear site is chosen or even existing sites expanded.

Ignoring present commissioned nuclear power stations, the Department of Energy's (DEN) 'central' energy forecasts imply about 135 GW installed nuclear capacity in the UK in the year 2025. If future sites are assumed to be 4 or 5 GW in size something like 30 sites would be required by 2025, or one new site less than every two years. Larger sites of, say, 8 GW would reduce the site acquisition rate, as would any further potential for building on existing sites. There is also a 15 GW 'policy gap' in the central forecasts for the year 2000 but there would be every prospect of this being met by an existing thermal reactor system. To many people, the DEN's forecasts seem too high; but the point here is to illustrate the implications for decision-making if they are correct.

The prospect of a biennial nuclear debate is particularly daunting because of two factors. If the urgency of energy decision-making is accepted, then the prospect of a filibuster against

nuclear development along the lines of the one that has developed in the United States becomes real. Such a filibuster might be attractive to anti-nuclear groups if it were not likely to generate widespread apathy about the issues through continuous repetition. Keeping the local inquiry as a possible context for energy debates should give no satisfaction to either the pro- or anti-nuclear camps.

How should an energy future be decided?

The choice of decision-making process for energy futures will affect the nature of that future. If the 'policy gap' can be met by some combination of combined heat and power schemes, solar installations, conservation, heat pumps and perhaps wind and wave power it seems unlikely that the existing *ad hoc* procedures for public participation need be altered. In particular, the local inquiry must stay to deal with what the local people want.

On the other hand, a coal-dominated scenario, however, would raise 'national' concerns, since the prospect of doubling coal output by 2025 would involve significant developments on the scale of the Selby and Belvoir coalfields. But nuclear dominated future demands the greatest departure from existing procedures: the belief that there are issues of a profound ethical nature is sufficient to justify this dichotomy.

If the local inquiry is not the most efficient context for making such decisions—although it must remain the proper context for a siting decision—how are the views of pressure groups and the public to be heard? Pressure groups and the public are not necessarily one and the same. The former tend to be genuinely concerned, but professional, elites with no direct or obvious mass support. Some believe they have this support, but most would admit that their role is an educative and persuasive one.

Sorry, for copyright reasons some images on this page may not be available online

The authors are researching into the implications of the Windscale Inquiry for future energy decisions in the UK. They wish to acknowledge a grant from the Energy Panel of the Social Science Research Council.

Ivor Nicholas

A pause in the debate at the Windscale inquiry

The public, on the other hand, has limited access to information about the issues involved in the nuclear debate. The answer may lie in information campaigns of the kind that have taken place in Austria, Denmark and Sweden where government funds are used to produce neutral documents conveying the facts as they might seem to a disinterested party, although the problems of defining a 'neutral fact' cannot be underestimated. Pressure groups may also be funded to produce their own literature and campaigns given that nuclear authorities already fund their own less explicit campaigns and their own defence at public inquiries. The alternative is to leave the debate to two professional elites—pressure groups and the establishment. This will happen anyway and the extra dimension of public participation may not materialise. Yet the public should at least have the opportunity of expressing its concern.

Once sufficiently informed, how are the people to convey their views to the elected? The normal processes for relay between constituent and parliamentary member seem redundant if no political party is to take up the cause. At the moment, energy futures do not divide along party lines. For extra-parliamentary debates the pressure groups are not on equal terms with their opponents since their funds are limited even if their expertise is often commendably high. If the debate is to be fair it would seem right to fund their activities on the basis of 'open government'. The dilemma arises in trying to define who constitutes a pressure group and how accountable they are to be in their use of funds. Failure to supply funds could be construed as a strategic mistake since a pressure group has no better excuse for resort to 'civil disobedience' campaigns than that all other channels have been closed to it. Yet if funds were supplied the pressure groups would be required to accept the outcome of any debate and the pro-nuclear camp is quick to point out that some pressure groups may never take 'no' for an answer. Equally, pro-nuclear pressure will not disappear if the debate goes against the establishment policy.

Another problem is whether any government information programme, on the lines of the Austrian, Swedish, and Danish models, should include a preliminary assessment of the project in question, perhaps taking the form of a cost-benefit study or an environmental impact statement? Apart from the problems of determining who will undertake such work, is any purpose served by such documents beyond what is achieved by the submissions of the various parties to the debate? The quality of submissions would vary

Sorry, for copyright reasons some images on this page may not be available online

David Swift

Mr. Justice Parker investigates at the Windscale works in Cumbria

widely, although aid to objecting parties could help. Nor do we escape the problem of putting the issues before the public if the impact statement or cost-benefit study is comprehensible only to the professionals. Yet there are real dangers in reducing the issues to the 'common sense' level, for it may be that some issues are just not reducible to common sense with the sacrifice of accuracy.

All this leaves the fundamental question still unanswered. Accepting the need for some extra-parliamentary debate involving pressure groups and at least a campaign to inform the public, what is the proper structure and form of that debate?

Issues not objectors

Basic principles of organisation seem clear. The debate should proceed by issue and not by some sequence of objectors, frequently overlapping and repeating views, as at Windscale. The adversarial context has its role if only because it permits cross examination and the extraction of information from parties to the debate, information that otherwise might not emerge. But whether or not the context needs to be a legal one is another matter. There is much to be said for the view that lawyers 'intervene' between experts, losing arguments for want of knowledge. Equally, if the adversarial context is a requirement then their training is partly fitted to the context. The quasi-judicial framework fails in other respects since it again favours the professional and discriminates against those who may not articulate their views in the same way.

The form of the institution responsible for the debate is also dependent on a policy decision. Is any decision about the nuclear future a once-and-for-all decision, or one that

we can expect to see debated again and again? The technological argument seems to favour the second option since a commitment to one fast reactor should not imply a necessary commitment to the second. Nonetheless, one can appreciate the fears of those who see the sequence as being irreversible in practice, a kind of nuclear future by stealth.

Opposition to nuclear power is not likely to go away because one national debate decides against it, any more than the reverse is the case. The institutional context should therefore be semi-permanent, capable of being re-constituted when the debate is renewed. No existing standing authority would seem to meet any of the requirements laid down for equal access by people of opposing views—as an example the Energy Commission contains no representative of pressure groups outside the fuel industries themselves, their trade unions and a few individuals.

A standing commission would seem to favour the Royal Commission format. Yet a Royal Commission has no formal link with Parliament: it can only advise and there is no requirement on the Minister to make any decision on its recommendations, whereas he is (theoretically, anyway) with the local inquiry. The Commission in question must therefore have a recommendatory power which obliges the relevant Minister to deliberate on its recommendations. Further, there should be no obstacle between the recommendation and parliamentary debate *before* a decision is made. Otherwise we shall witness the growth of a decision-making framework which by-passes Parliament. An Inquiry Commission empowered to sit whenever the debate seems necessary will be costly. It may, however, be the only form of institution that meets the requirements for efficiency in an imperfect world.

Arguably, all this concern comes too late. North Sea oil will be very short-lived and the energy demands of the 1990s must be met by technologies about which we must decide now. A decision-making procedure that postpones decisions could be costly in terms of employment and income. Equally, it would be a naive industry that did not plan for such delays and perhaps there is time for a wholesale debate. Probably, there is no alternative. Failure to establish the requisite institution now leaves us with the *ad hoc* procedures we already have and we have shown that they can readily generate the same problem of delay.

Failure to allow any debate through either existing or new institutions will merely lead some pressure groups to adopt less democratic forms of action that might threaten the very basis of decision-making itself. □

UK report recommends nuclear reprocessing

THE long-awaited 'Windscale report', the result of a 100-day inquiry into a planning application for a nuclear fuel reprocessing plant, was published this week by a subterfuge on the part of the minister responsible, Mr Peter Shore, Secretary of State for the Environment. The report recommends that outline planning permission should be granted "without delay" to British Nuclear Fuels Limited (BNFL) to build a new thermal oxide reprocessing plant (THORP) at Windscale, Cumbria. Mr Shore, legally bound to make a decision on the issue without further consultation, but wishing to involve parliament in the debate, has decided not to grant planning permission, but to seek a 'Special Development Order' through parliament which would reverse his decision and allow the plant to be built. Mr Shore is in favour of the plant; the arcane procedure is the result of the lack of an effective mechanism for reaching democratic technical decisions in the UK—a problem in which the UK is not alone.

BNFL began lobbying MPs some months ago in the expectation that a debate was likely. A substantial and vociferous lobby of MPs is against THORP, but BNFL are "99% sure" that THORP will survive the parliamentary process. Costing £600 million at current prices and designed to reprocess 600 tonnes of British and foreign spent oxide fuel a year by the PUREX process (see below), THORP will take 10 years to build once authorised.

The report, the lonely work of the Inspector, Mr Justice Parker, has been received by BNFL as a "complete vindication" of their proposals, and by Mr Shore as "cogent and persuasive". Friends of the Earth, on the other hand, "found it hard to credit the extent to which he [Mr Justice Parker] has overlooked or misunderstood key aspects of the argument". FOE single out the passages on waste management, energy economics and foreign policy for special attack. "We can only assume that the pressure to produce the report quickly left insufficient time to assimilate the evidence" said FOE.

Mr Parker's support for THORP is based on 12 principal arguments, which are as follows:

1. Stocks of spent fuel from AGRs (advanced gas-cooled reactors) presently existing and under construction will, unless reprocessed, continue to build up and will have to be stored until disposed of in some manner.
2. It is necessary to keep the nuclear industry alive and able to expand should expansion be required.
3. Keeping the industry alive will involve further reactors being constructed and further quantities of spent fuel arising.
4. All the spent fuel stored will contain plutonium. The inventory of plutonium will therefore continue to increase for as long as reprocessing is delayed.
5. Prolonged storage of spent fuel would involve the development of new storage methods, which would be a costly and lengthy process.

6. To store increasing quantities of spent fuel would only be sensible if it was ultimately likely to be decided to dispose of the spent fuel without reprocessing.

7. Such a decision appears to be unlikely and not in our best interests or in those of future generations; it would commit future generations to the risk of escape of more plutonium than is necessary; and the risk would be greater since the spent fuel is likely to be more vulnerable to leaching by water than solidified highly active waste.

8. If reprocessing is going to take place at some time then it is preferable to start without delay, to gain experience of the process and its dangers while amounts of fuel to be reprocessed are small.

9. The risks from the emissions involved in reprocessing are likely to be very small and, if reprocessing is to be designed in to THORP if they proved correct.

10. The risks of accident will, if reprocessing is to take place at some time, will in any event occur at some time. Evidence that current estimates are seriously wrong "did not appear to be convincing" wrote Mr Parker but any new estimates would ultimately have to be time, also have to be incurred, at some time. At the present they are likely to be containable within tolerable levels. If reprocessing were to begin suddenly on a large scale after a delay, risks would be greater.

11. The risks from terrorism are not significant.

12. The risks arising from transport would be no greater than at present.

Robert Walgate

New reprocessing technique promises diversion safeguards

BRITISH and American nuclear scientists have designed a new system for reprocessing spent reactor fuel which, it is claimed, makes the diversion of nuclear fuel for military purposes virtually impossible by ensuring that pure plutonium is not accessible at any part of the cycle.

The new system was announced jointly by scientists from the United Kingdom Atomic Energy Authority (UKAEA) and the Electric Power Research Institute (EPRI) of Palo Alto, California at a conference on energy technology held last week in Washington.

Developed in direct response to President Carter's concern that the worldwide expansion of nuclear power

could lead to the increased proliferation of nuclear weapons, it is hoped by the nuclear community that the new process will calm many fears and open up the way for the 'safe' development of fast breeder reactors.

In contrast to conventional reprocessing techniques, whose prime aim is to separate pure plutonium and pure uranium from spent reactor fuel, the new system retains the plutonium mixed with both uranium and fission products making it lethally radioactive. In addition, the technology of the reprocessing process has been designed in such a way that even if a large force took over the plant, it would be unable, without making major and time-consuming modifications, to divert the

process into producing pure plutonium that could be easily handled.

Speaking in Washington on Monday, Dr Walter Marshall, Deputy Chairman of the UKAEA and until recently chief scientist at the UK's Department of Energy, said that the UKAEA and the EPRI shared a belief that once a fast breeder reactor cycle had been developed, then it could be made proliferation proof. "You can make it so difficult to steal the plutonium from the cycle that you can virtually forget about it" Dr Marshall said.

In a paper to the conference prepared jointly with Dr Marshall, Dr Chauncy Starr, President of EPRI, said that recent concern over, for example, the possible theft of plutonium by terrorist groups made it important for the future guarantee of world energy supplies to develop a joint reactor and reprocessing system that was diversion-proof, in the sense that the difficulty

of extracting weapons material was roughly equal to or greater than that of doing so by independently processing long-stored fuel.

The fact that, in contrast to thermal reactors, fast breeder reactors were relatively insensitive to fission products in their fuel meant that it was possible to conceive of "a hypothetical idealised breeder fuel cycle which at all points has a plutonium-uranium mixture that does not exceed the 15 to 20% plutonium necessary for fresh fuel, that is only partially decontaminated from fission products, and is therefore highly inaccessible" Dr Starr said.

The net result of such a reprocessing system combined with a fast breeder reactor would be the creation of a diversion-proof nuclear power capacity that would effectively remove such nuclear power systems as a potential resource for weapons proliferation.

So far the new system, which is being submitted to the International Nuclear Fuel Cycle Evaluation Study (INFCE) set up last year at the suggestion of President Carter has received a cautiously optimistic welcome from the administration, keen to develop a politically-acceptable fast breeder nuclear reactor programme and reprocessing policy.

A spokesman for the US Secretary of Energy, Dr James R. Schlesinger, said that the concept described by the scientists would receive serious consideration because it fitted with the goals of the federal reactor research programme.

However, environmentalists still have their doubts. In a joint statement, two Washington-based environmentalists groups, New Directions and the Natural Resources Defense Council, criticised the new proposal for a failure to set

high enough standards of safety, saying that the new technique could be used as a cover-up for the production of nuclear weapons.

The relative desirability of fast breeders over thermal reactors—even taking proliferation dangers into account—had been emphasised by Dr Marshall four days prior to the Washington meeting when he gave the Graham Young Memorial Lecture at Glasgow University in Scotland. Dr Marshall said that a policy of using thermal reactors alone in the once-through cycle was not a satisfactory non-proliferation policy since every fuel storage facility became, in essence, a 'plutonium mine'. Furthermore the net rate or production of plutonium by fast breeders was potentially lower than the production of plutonium as waste by thermal reactors.

David Dickson

How 'CIVEX' and 'PUREX' differ

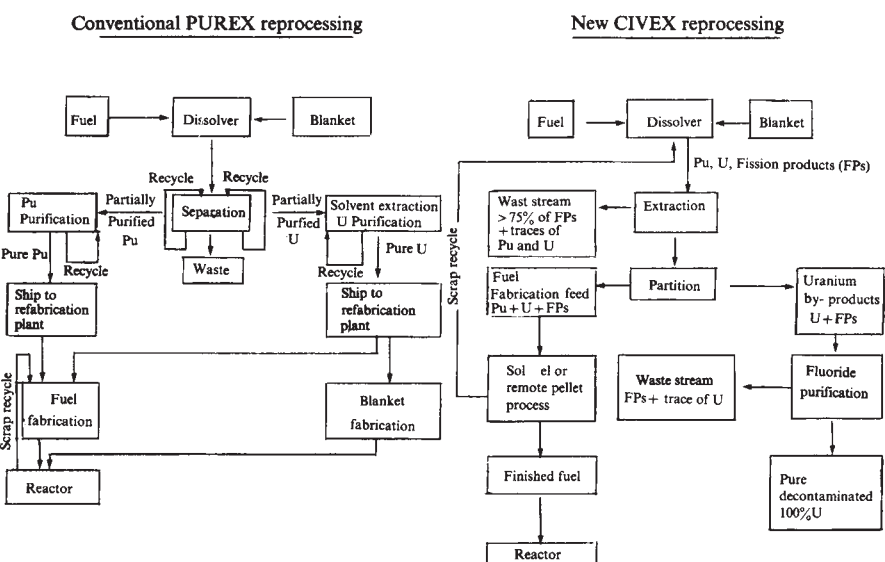
DETAILS of the new reprocessing system—described as 'CIVEX' to emphasise its civilian character and distinguish it from the conventional 'PUREX' reprocessing process—were described to the Washington conference by Dr Milton Levenson, Director of EPRI's nuclear power division.

According to Dr Levenson, seven criteria were necessary to help define a diversion-proof reprocessing system. These were:

- No pure plutonium in storage
- No pure plutonium at any intermediate point
- No way to produce pure plutonium by simple process adjustment
- No way to produce pure plutonium without equipment modifications
- No way to carry out equipment modification with facilities and components normally on site
- No way to carry out the required equipment modifications without plant decontamination or entry into extremely high radiation fields
- Length of time required for successful diversion such that adequate time is available for national and/or international response.

Dr Levenson said that a number of steps in the conventional PUREX process violated at least one of these criteria. Such unacceptable steps included the shipment of pure plutonium from a reprocessing plant to a refabrication plant, the final plutonium purification cycles, and the provision for recycling material that was not highly purified.

The first steps of the solvent



extraction operation prior to the scrubbing of fission products from the plutonium/uranium streams did meet the criteria, and would be retained in the CIVEX process. However, the process was new in two ways, Dr Levenson said.

Firstly the chemical steps used for uranium purification use a fluoride purification process which is effective for purifying uranium but not for plutonium. The excess uranium which is to be recycled for blanket fabrication is collected for subsequent purification by a fluoride volatility process using bromine trifluoride or low temperature fluoridation to produce UF_6 and subsequent purification by distillation or sorption-desorption.

Furthermore the plant equipment and layout would be such that there is no way to change the mode of

operation to produce plutonium. In the PUREX process, whose objective is to make the purest possible uranium and plutonium, free of radioactive wastes, equipment is provided to permit the recycling and decontamination of any material carrying radioactive impurities. Such equipment is not present in the CIVEX plant, and concrete process cells are constructed of such a size as to make it impractical or impossible to install any.

Further modifications in the CIVEX design include the absence of a separate scrap recovery facility, and the fact that the primary product stream is taken directly through a remote fabrication operation, using the sol-gel method of making oxide or remote application of the more conventional oxide process through to finished fuel.

Russian science magazine dates from before the revolution

PRIRODA, the Russian popular science journal, appeared in its 750th issue last month, with the original *art nouveau* cover of 1912, and containing reprints of significant articles published in the last 66 years. Any cover from the early years of the century must inevitably have an aura of the antique, but the first *Priroda*, with its antiquated spellings and obsolete letters comes as something of a shock—the very orthography is at once a living reminder that this is a journal from before the Revolution.

Priroda was founded in 1912, the same year in which *Pravda*, albeit illegally and abroad, first saw the light of day. *Priroda* was founded as a popular science journal, under the auspices of the "Man and Universe" Society, and its pre-Revolutionary editors included Vladimir A. Vagner (biologist and physiologist), Lev V. Pisarzhevskii (chemist), Lev A. Tarasevich (microbiologist) and Nikolai K. Kol'tsov (biochemist). From this era, the jubilee issue reprints Il'ya I. Mechnikov's 1915 address to the Pasteur Institute, and a fascinating article by Maksim Gor'kii on "Science and democracy" which appeared in May 1917, during the heady political turmoil between the February and October Revolutions.

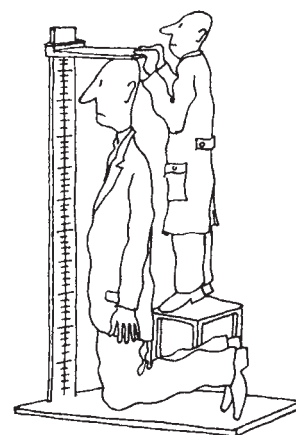
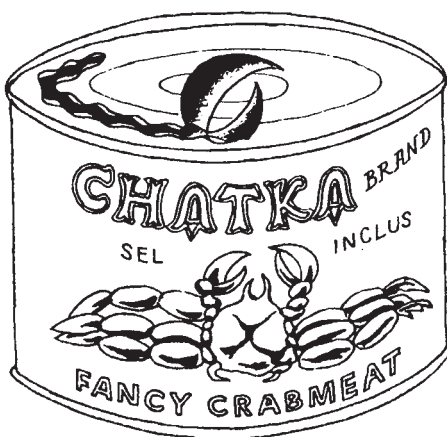
Revolution, however exciting, does not favour the publication of science, although *Priroda* managed to struggle through, albeit with occasional double issues until 1920. The following year, it was refounded as the organ of the "Commission for the Study of Natural Productive forces of the Academy of Sciences of the USSR" and the place of publication transferred from Moscow to Petrograd. In 1930 it came under the direct auspices of the Academy. During these years, the nature of the journal changed, becoming increasingly specialised, although some papers, such as Pavlov on "The activity of the higher nervous system" (1932), Vernadskii on the evolution of species and Artsimovich and Alikhan'yan on "the problems of the physics of the atomic nucleus" (1941) still seem aimed at the interested amateur rather than the specialist. An interesting reprint from this time is "The problem of the genesis of cultivated plants in our present state of knowledge", by Nikolai I. Vavilov, who was later to be discredited by the Lysenkoists and was only fully (post-humously) rehabilitated in 1971.

In 1951, a special resolution of the Academy restored *Priroda* to its original purpose, the popularisation of science, and the specialist articles were

banished to more learned forums. At the same time, the editorial offices were moved back to Moscow. The format of *Priroda*, with prominent scientists, often Academicians, writing for the popular market survey articles on the achievements of Soviet and foreign science became essentially that we know today. With the passing of Stalinism, humour took a lighter touch—one recalls in particular the cartoon of the female biochemist earnestly knitting a double helix, while at the same time often subtly critical of official policy or negligence (the lab. assistant

December 1998 and *Priroda* 1000. (Articles 'reviewed' include "The radiation of blue holes in the pink part of the spectrum", "The proposed international alphysical dictionary", a declaration from the Academy that they will not consider articles on telepathy "since the telepathic link has been conclusively proved", reports of traces of organic compounds in moon rock, and Dirac monopoles whose vibrations reproduce the "old-time love-song" *Ochi cherniya!*)

Turning from fantasy to fact, and congratulating *Priroda* on its enduring standards and impressive survival and equally impressive circulation figure (85,000), one must not fail to remark one minor



Priroda style humour: the saving mutation (left): science always finds a way out (right)

throwing tin-cans and cigarette ash into the fish-tank "to create a natural environment"). Humour is not lacking from the jubilee issue, taking the form of the regular round-up of literature—but in this case looking forward to

detail. The 1912 cover gives the price of a single issue as 50 kopeks. Two world wars, two revolutions and 66 years later the cost of a single issue remains—50 kopeks.

Vera Rich

Czechs and Soviets meet in record space flights

CZECHOSLOVAKIA has won the "little space race" for the third nation to put a citizen into orbit. The Czechs were odds-on favourites, although Poland and East Germany are confidently expected to 'place' later this year. The launch of 'Pilot Cosmonaut' of the Czechoslovak Socialist Republic' Vladimir Remek, as one of the crew of *Soyuz-28* is not, of course, unexpected; the *Interkosmos* programme envisaged the participation of Comecon cosmonauts, each with a Soviet partner, in *Soyuz* flights starting this year, and it has been rumoured for some weeks that these "international" flights would dock with the *Salyut* station. The only contention was which of the eight 'junior partners' in the alliance would be first.

Czechoslovakia has from the beginning been a major contributor to

Interkosmos, both in technical expertise and the provision of tracking facilities. Nevertheless, the Czech press played down this aspect, maintaining (*Rude Pravo*) that the joint flight was an example of Soviet selflessness and the principles of proletarian and specialist internationalism, and stressing (*Zemedske Noviny*) that the Soviet Union bore about 95% of the cost of the *Interkosmos* programme—a claim that can neither be confirmed nor refuted, in view of Soviet reluctance to publish estimates for the space programme.

The new 'international' link-up was certainly the occasion for considerable euphoria—both at the control centre 'somewhere near Kalingrad' and in orbit. Indeed, the celebrations aloft seem to have exceeded expectations, since, according to the flight director

A. Eliseev, the four cosmonauts "despite a great deal of persuasion from control" had not quietened down until late.

In spite of the excitement of this second linkage, the experimental programme of the mission is continuing to plan. The new crew are involved in an experiment with chlorella and have placed a capsule in the 'Splav' alloying chamber, which is designed to produce alloys of, for example, aluminium-tungsten and molybdenum-tungsten, which would be impossible except in low or zero gravity conditions. (The particular 'Splav' experiment in which Remek is involved has been named 'Morava' in honour of Czechoslovakia).

However, according to Yakov Ziman of the Soviet Institute of Space Research, the space programme for 1976-80 (including the Comecon *Interkosmos* programme) is particularly concerned with the earth. Not surprisingly, the over-all plans for Salyut-6 have included a considerable geophysical programme which has been going forward steadily throughout the mission. Whereas the Soyuz-22 photographs were "purely scientific", only one-tenth of the Salyut-6 survey is devoted to pure research; the rest are tailored to the requirements of various Ministries, government departments and design organisations.

Much of the data brought back by the crew which visited Salyut at the end of January was at once processed and transmitted to these 'clients', including various organisations involved in opening up the route of the Baikal-Amur Mainline railway, and the constructors of the new hydroelectric installations in the highly-seismic vicinity of the Nurek dam.

Other photographic surveys include the monitoring of ocean currents (including oil slicks), and a special study of the glaciers and snow cover of the mountains of Soviet Central Asia. Space photographs are used to monitor glacier movements—a recent TASS announcement claims that over 30,000 "pulsating" glaciers have been discovered in the Pamirs using space photography.

Grechko is apparently a fairly talented artist in the 'space' medium of coloured pencils and felt-tipped pens. This gift is being used in the observation of ocean currents and the aurora borealis. In the case of the ocean observations, there seem to have been some initial difficulties in recording the currents—the colours observed went beyond the scale provided by the oceanologists. Special attention in all such visual observations is being paid to a phenomenon noted by the crew of Soyuz-25—under certain circumstances cloud cover can act as a "gigantic lens" so that waves, for example,

appear considerably magnified.

Other phenomena sketched by Grechko include Venus rises and sunrise over a clear horizon. Some of these sketches were brought back by *Soyuz-26* and are being studied at the Leningrad Optical Institute. The aurora borealis appearances sketched by Grechko coincided with observations of noctilucent clouds, although, according to Professor Aleksandr I. Lazarev of the Leningrad Optical Institute, it was difficult to establish any link between the two. These were the first space observations of noc-

tilucent clouds in the southern hemisphere, although they had been observed in the northern hemisphere from Salyut-4.

By combining Grechko's observations with data from meteorological rockets launched from the Antarctic *Molodezhnaya* station it has been deduced that noctilucent clouds have a multi-layer structure, occurring at local temperature minima. A three-layer structure is postulated, distributed at temperatures of -130 to -150 °C.

Vera Rich

Voyager scientific platform jammed

VOYAGER 1, the spacecraft launched last summer by the National Aeronautics and Space Administration as part of a joint project to Jupiter and Saturn, has developed an equipment fault which poses a major threat to the scientific success of the mission.

The problem has been caused by a malfunctioning of the scan platform (shaded on diagram) which contains most of the major scientific equipment to be used on the mission, including two television cameras, a cosmic ray detector, a photopolarimeter and ultraviolet and infrared spectrometers.

The platform is designed to rotate along two separate axes—an azimuth axis and an elevation axis—in such a way that the equipment can be pointed at any desired object either in the sky or on the surface of the two planets.

During calibration tests two weeks ago, the movement of the platform on its azimuth axis slowed down to such an extent that it was unable to complete a particular manoeuvre in the one-hour's computer time allotted, and thus failed to reach the required position.

Further commands issued from earth met with the same response. And al-

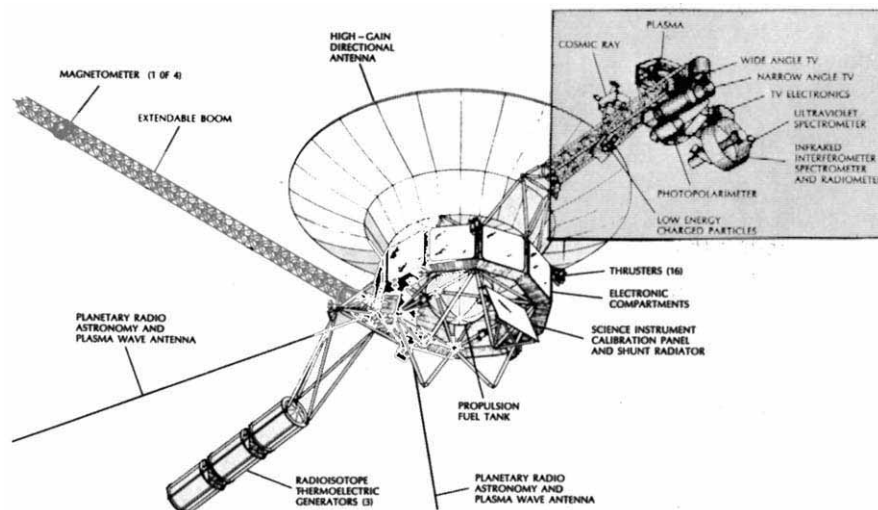
though it was initially suggested that the problem was due to an electronic failure in the spacecraft's computer system, it is now thought that the jamming of the platform is due to a mechanical cause.

Scientists at NASA's Jet Propulsion Laboratory in Pasadena, from which the Voyager mission is being directed, are making a hurried analysis of the spacecraft's pre-flight test data, to see if they can find any anomalies in performance which were previous overlooked.

Their intention is to produce a set of diagnostic commands which can be transmitted to the spacecraft. Voyager's reaction to these commands will, it is hoped, enable a fault to be analysed and corrected.

Dr Raymond L. Heacock, programme manager for the Voyager mission, said last week that even if it was impossible to rectify the fault, it would still be possible to get the equipment pointing in desired directions by manoeuvring the whole spacecraft. Thus scientific data could still be transmitted back to earth.

However such a procedure would



involve a more complex sequence of manoeuvres than initially planned, thus consuming greater quantities of fuel. In addition, the extra time needed to direct the equipment towards a particular target would severely limit the amount of data received.

Voyager 1 was launched last September, and is now over 200 million miles from earth. It is expected to start taking photographs of Jupiter and making spectral scans of the hydrogen

cloud surrounding the planet in December, 80 days before reaching the planet.

In late February 1979, eight days from Jupiter, it will begin coverage of the entire planet with its wide-angle camera, while the atmospheric infrared and ultraviolet spectrometers and the photopolarimeters will obtain data on atmospheric composition, temperature variation in the atmosphere, and the possibility of solid particles in the

clouds. Observation are, at present, planned to continue at least until April 1979.

A technical problem involving a jammed photopolarimeter filter wheel on board the sister spacecraft Voyager 2, which was launched 12 days earlier than Voyager 1 but on a slower trajectory, and will not reach Jupiter until July 1979, has already been successfully remedied.

David Dickson

Cassava may lead to mental retardation

CASSAVA is the world's seventh most important foodcrop, following the major cereals, potatoes, and yams. Production is in the neighbourhood of 100 million tons, and it is likely to double by the beginning of next century. A major advantage of cassava is that it can grow on poor soil with little rain; it is the principal source of carbohydrates for about 300 million people, most of them in the developing countries of the tropics.

This background underlines the importance of the findings reported by a team of Belgian researchers, headed by Dr André-Marie Ermans of the Department of Radioisotopes at the Saint-Pierre Hospital, University of Brussels. In a study of the population of Idjwi Island on Lake Kivu, Zaire, they have shown that a steady diet of cassava inhibits iodine uptake by the thyroid gland. When iodine supply is marginal, this can cause endemic goiter, cretinism, and mental retardation.

The Belgian researchers are now following up their study with a campaign in the Ubangi region, in the north-eastern part of Zaire, where goiter is endemic in a population of about one million. They have found that 60–70% of the inhabitants have endemic goiter, and 1–10% are affected with cretinism. In addition an unknown number of people suffer from varying degrees of mental retardation. The objective of the campaign, financed by the Belgian government, the Zaire Institute for Scientific Research, and Canada's International Development Research Centre, is to eradicate goiter and cretinism in the Ubangi region. Some 300 000 people have already been "vaccinated" by receiving intramuscular injections of an iodine suspension in oil, that diffuses into the organism over a period of three to seven years; 700,000 more injections are planned.

Congenital hypothyroidism, with its sequels, notably cretinism, is one of the most widespread diseases in the tropics. In the past ten years or so, epidemiological studies in Africa, South America and Asia have revealed that more than 200 million people may be

affected by goiter. How much cassava, as a staple food, may contribute to this is yet to be determined.

In the past great emphasis has been laid on research to improve cassava



Cassava grinder

productivity and utilisation, but the findings of the Belgian team point to new, and imperative, avenues of research: the prevention of cassava-mediated hypothyroidism, better ways of de-toxifying the tuber before it is consumed, and the development of new lines that do not contain the chemicals responsible for this form of toxicity. (Cassava contains cyanogenic gluco-

sides; when ingested, these glucosides are detoxified, yielding thiocyanate as a by-product which inhibits iodine uptake by the thyroid.)

These findings, of major importance to developing countries in the tropics, are not without implications for industrialised countries. In Central Europe and in regions along the Mediterranean, where there is limited supply of iodine in food, other vegetables, such as cabbage, may have a similar effect.

François Delange, a Belgian paediatrician, who is participating in the Zaire programme, has carried out a study in cooperation with Sicilian physicians, and found that high thiocyanate levels are associated with goiter on the island. It is known that in Belgium the iodine content of food is rather low, and tests carried out on 1,800 newborn in Brussels have shown that 14 of them had thyroid insufficiency. If this is not corrected at an early stage, some of these children may become mentally retarded.

The Belgian version of screening for congenital hypothyroidism has certain advantages. It relies on the measurement of the pituitary hormone TSH (thyroid stimulating hormone). A high level indicates transient or permanent thyroid insufficiency. The advantages of the test is that it can be made on a single drop of dried blood, that it can be automated, and is relatively inexpensive (about \$3). It is likely that, as of the end of the year, it will be administered routinely to all newborn children in Belgium.

Alexandre Dorozynski

Genetically engineered bacterium patented

The first patent for a genetically engineered microorganism has been granted by the US Court of Patent Appeals to General Electric for a bacterium which degrades crude oil more completely than any bacterium found in nature. A strain of *Pseudomonas* that can degrade about 60% of crude oil has been developed in General Electric's New York Laboratories by one of their biologists, Dr A. Chakrabarty.

Many strains of *Pseudomonas* naturally contain small loops of DNA coding for the breakdown of many of the complex organic compounds making up crude oil, but each strain by itself can only degrade a few of the compounds. Over the past few years, Dr Chakrabarty has constructed strains of *Pseudomonas* containing several of these DNA plasmids, thus extending the range of compounds that this bacterium can attack. □

correspondence

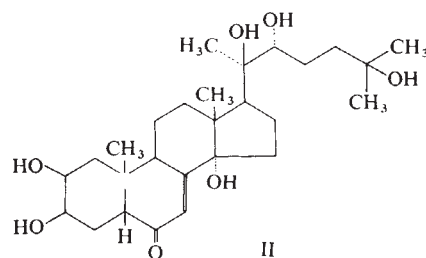
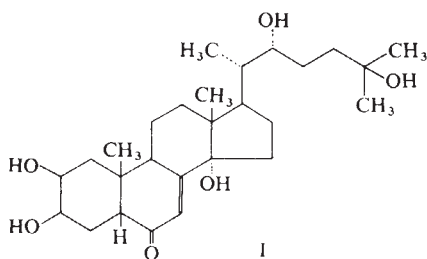
Ecdysteroids: a new generic term

SIR,—Following the isolation from plant and insect sources of steroids with molting hormone activity in insects, the term ecdysone originally given to one defined compound (I) was used by many authors as a generic term for all substances with molting hormone activity. We feel that it is highly undesirable to use the same name as a generic term and simultaneously as the name for a specific compound. We therefore propose the term 'ecdysteroid(s)' as a generic term for all compounds structurally related to ecdysone. The name is coined in analogy to corticosteroids, a well known generic term comprising steroid hormones from the adrenal cortex.

Since many steroids with molting hormone activity have been isolated from plant and animal sources, and more will be isolated in the near future, a proper generic name is highly desirable. We urge adoption of the generic name ecdysteroids.

If a subdivision of this class of compounds to source is required, the names 'zoecdysteroids' and 'phytoecdysteroids' may be used. The generic term 'phytoecdysones' should be abandoned.

The name 'ecdysone' should only be used for the molting hormone isolated from insect sources by Butenandt and Karlson¹ represented by formula I, that



is 2 β , 3 β , 14 α , 22R, 25-pentahydroxy-5 β -cholest-7-en-6-one.

Compound (II) 2 β , 3 β , 14 α , 20R, 22R, 25-hexahydroxy-5 β -cholest-7-en-6-one or 20-hydroxyecdysone, has been isolated from natural sources independently several times: as β -ecdysone^{2,3}, as

ecdysterone⁴, as crustecdystone⁵, as isoinokosterone⁶, and⁷ as polypodine A. We propose to use the term '20-hydroxyecdysone' as the name for this biologically important molecule. In the physiology of insects and crustaceans, the terms 'ecdysterone' and 'crustecdysone' may be used. The isolation from crustacean material can claim priority to all other reisolations, since the substance described as β -ecdysone was not pure 20-hydroxyecdysone. But the name '20-hydroxyecdysone' stresses the close relationship to ecdysone; this, we feel, is desirable.

Yours faithfully,

T. W. GOODWIN
University of Liverpool, UK

D. H. S. HORN
C.S.I.R.O. Melbourne, Australia

P. KARLSON
J. KOOLMAN
University of Marburg, West Germany

K. NAKANISHI
Columbia University, New York, USA

W. E. ROBBINS
U.S. Department of Agriculture,
Beltsville, Maryland, USA

J. B. SIDDALL
Zoecon Corporation, California, USA

T. TAKEMOTO
Tokushima Bunri University, Japan

1. Butenandt, A. & Karlson, P., *Z. Naturforsch.* **9b** 389–391 (1954).
2. Karlson, P., *Congress Internat. de Biochimie Bruxelles Abstract* 12–3 (1955); *Ann. Sci. Natur. Zool.* **18**, 125–137 (1956).
3. Hocks, P., Schulz, G., Watzke, E. & Karlson, P., *Naturwiss.* **54**, 44–45 (1967).
4. Hoffmeister, H., *Angew. Chemie* **78**, 269–270 (1966).
5. Hampshire, F. & Horn, D. H. S., *Chem. Commun.* 37–38 (1966).
6. Takemoto, T., Ogawa, S. & Nishimoto, N., *Yakugaku Zasshi* **87**, 325 (1967).
7. Jizba, J., Herout, V. & Sorm, F., *Tetrahedron Letters* 1689–1699 (1967).

Catastrophe theory

SIR,—Zahler's recent letter (2 February, page 401) contains a number of misapprehensions and errors. I have no wish to prolong the correspondence on the applications of catastrophe theory so I will confine myself to correcting his misunderstanding of what I said in my letter (22/29 December, page 658) about Kozak and Benham's model for protein denaturation.

My letter does not give "another reason why catastrophe theory should not apply to denaturation". It explains why protein denaturation cannot conform to the cusp catastrophe, as Kozak and Benham agree (private communication). In their later paper (*J. Theor. Bio.* **63**, page 125), Kozak and Benham use catastrophe theory to discuss the continuous transition of protein denaturation by means of a distribution of catastrophe potential functions. Zahler's observation that their use of the Maxwell convention "removes the cusp and thus catastrophe theory from the analysis" is correct to the extent that under the Maxwell convention the catastrophe set is a line and not a cusp. But this is irrelevant and the latter part of the quotation incorrect because the elementary catastrophes are described geometrically and do not depend on the

choice of convention for determining the governing minimum of the potential function.

Yours faithfully,

MAURICE DODSON
University of York, UK

Sunflower's blooming floscule is a compass

SIR,—Two letters (8 September, page 102 and 15 December, page 556) correctly describe the behaviour of the top (and the bud) of growing sunflowers *Helianthus annuus*. Such movements are nutational ones. (The picture with the second letter is unfortunately incorrect: they should be buds.)

I should like to add the following. When the basis of the bud (the part between the stem and the bud) becomes steep instead of smooth, these nutational movements get smaller and smaller in diapason and point less and less to the West. Finally, before blooming the floscules stop their movements completely and (for them) forever. Most parts of blooming floscules are orientated to the sunrise. In the mornings they look at the sun and so give an impression of unanimous movements after the sun. This is the only time of day when this is true.

That is why blooming (and faded) floscules of the sunflower are living compasses (however not too exact): west is behind, north to the left, and south to the right. This is true for open country. If the sunflowers are shaded by a wall from the east, their blooming floscules look (permanently) to the west side. Perhaps they need UV for seed formation, the intensity of which is the most at sunrise, least at midday and intermediate at sunset.

Another question also seems interesting. In different languages the sunflower has different meanings, for example:

- "A flower (rose etc.) of the sun" in Latin, English, Swedish, Dutch, German, Rumanian, Armenian, Buryat, Mongolian, Kalmyk, Altaian, Latvian, Estonian, Finnish, Afghan, Burmese, Czech, Polish, Ukrainian, Byelorussian, Danish, Greek;
- "Turning after the sun", "Looking at the sun" in Spanish, Portuguese, Italian, French, Bulgarian, Hungarian, Moldovian, Uzbek, Bashkir, Serbian, Azerbaijani, Kirghiz, Turkmen, Japanese, Chinese, Viet-Namese;
- "Under the sun" in Russian (and one of the Serbian dialects);
- "Moon flowers", "Looking at the moon" in Turkish, Karakalpak.

Unfortunately, I do not know how to translate the names of the sunflowers from (American) Indian languages: Navaho—"ntigiliitahoh", Hopi—"tcego a gaw u" and Keres—"hi cti".

Yours faithfully,

G. G. POLIKARPOV
Principality of Monaco

news and views

DNA conservation and speciation: adaptive or accidental?

from Gabriel Dover

In the absence of a fundamental theorem of the genetics of speciation it is understandably tempting to relate any aspect of the structure of the genome to the processes involved in the phenotypic divergence of populations and species. The recent interesting conjecture of a possible causal relationship between the abundance and distribution of highly conserved sequences of satellite DNA and speciation in populations occupying the same geographical area (sympatric speciation) (Fry & Salser *Cell* **12**, 1069; 1977) is an extension of a long line of speculation on the evolutionary role of chromosomal differences in speciation that goes back to the classical studies of Dobzhansky in *Drosophila pseudoobscura*. Dobzhansky and his colleagues considered the association of particular alleles with particular chromosomal inversions in polymorphic populations to be due to the suppression of meiotic recombination in inversion heterozygotes and that such a restriction on exchange maintains presumed coadapted sets of alleles in strong linkage disequilibrium. Salser and his group suggest that a possible failure of meiotic pairing would result in the effective genetic isolation of populations differing in amounts of the same satellite DNA and thus facilitate a process of sympatric speciation. Unlike the studies on chromosomal inversions, however, the appeal to meiosis as the process whose proper functioning is disturbed in presumed structural heterozygotes of satellite DNA is purely speculative.

The appeal, nevertheless is made against a background of hard data that shows that not only do substantial portions of the genomes of higher organisms display a remarkable sequence conservation over long periods of evolutionary time, but also that the same conserved sequences are strikingly similar in distantly related species.

The question to be asked is whether the present-day distribution of con-

served sequences is due to random processes of generation or maintenance, or to a process of selection on some as yet undetermined function. In the absence of any information on these processes any hypothesis proposing that satellite sequences are involved in speciation either through their role in meiosis or in chromosomal rearrangement (Hatch *et al.*, *Chromosoma* **58**, 155; 1976) need not of necessity invoke natural selection.

The argument of Fry and Salser for a selective process is based on the analysis of satellite sequences in related species of kangaroo-rats and their comparison with those of distantly related rodents. One of the three satellite DNAs (HS- α) of the kangaroo-rat (*Dipodomys ordii*) occurs in at least 12 sequence variants, present in different amounts, which can all be related by one or two base substitutions to the predominant (about 25% of the total) 6-base sequence 5' TTAGGG 3', (Fry & Salser *Cell op. cit.*). None of these sequences is related to the sequences of the other two satellites either in the same species or in other related kangaroo-rat species (Salser *et al. Fed. Proc.* **35**, 23, 1976); showing that in this group there is no common ancestry for the different satellite types. Surprisingly, however, the major 6-base sequence is identical to the repeating unit of guinea-pig α -satellite (Southern *Nature* **227**, 794; 1970); furthermore finger-print analysis of the α -satellites of the pocket-gopher and the antelope ground squirrel shows that there is every reason to believe that the major and all the minor sequence-variants of HS- α satellite are similarly represented in these species.

These four species belong to three separate suborders of rodents that diverged 40–50 Myr ago. The remarkable extent of sequence homology rules out chance convergence to the same sequences in different suborders. The alternative suggestion of Salser and his colleagues is that all species within a group share a 'library' of sequences whose members may be arbitrarily amplified in different species but are

selectively maintained by the acquisition of unspecified functions. Members of the 'library' need not necessarily be related in sequence. In kangaroo-rats for example, the members of the 'library' are clearly not related; neither are the four satellite DNAs in man which are also distributed to different extents in three other primate species: a distribution that might be due to amplification of selected members of a primate 'library' (Gosden *et al. Chromosoma* **63**, 253; 1977). In other species groups, however, the sequences of the different satellites found within and between species, are related; in the *D. virilis* species group for example (Gall & Atherton, *J. molec. Biol.* **85**, 633; 1974), and in three of the four satellites of *D. melanogaster* (Brutlag *et al.*, *Cold Spring Harbor Symp. quant. Biol.* **42**, in the press), and sibling species, *D. simulans* (Peacock *et al.*, *Cold Spring Harbor Symp. quant. Biol.* **42**, in the press) indicating that in these instances they might have arisen from members of a 'library' of related sequences.

Unusual, and as yet inexplicable, conservation of sequences other than the highly repeated satellite DNA, for example, middle-repetitive and single copy, have also been reported recently by Flavell in related species of Gramineae (Flavell *et al.*, *Chromosoma* **63**, 205; 1977); by Macgregor, in salamanders (Mizuno *et al.*, *Chromosoma* **58**, 1976); and by Britten's and Davidson's group at CalTech, in sea urchins (Angerer *et al.*, *Chromosoma* **56**, 213; 1976).

Is it possible that one mechanism of sequence conservation, the consequence of some unusual behaviour of replicating DNA or replicating chromosomes underlies these diverse phenomena? Fry and Salser conclude that the various models proposed to account for the random generation and maintenance of sequence homogeneity (of which George Smith's model of repeated non-reciprocal exchange of sister chromatids is the front runner (*Science* **191**, 528; 1976; TIBS, N48, Feb. 1978)) are inadequate to explain

Gabriel Dover is a lecturer in the Department of Genetics, University of Cambridge, and a Fellow of King's College.

the preservation of similar sequence variants in species of rodents separated by 40 Myr, and instead they propose that selection has operated on some as yet unknown function of satellite DNAs.

They base their arguments on the rate of evolutionary divergence ($\sim 3\%$) in mouse satellite DNA that was assessed by Southern (*J. molec. Biol.* **94**, 51; 1975) from the relative amounts of monomers and multimers of fragments of DNA produced after a complete digestion of mouse satellite with restriction endonuclease *Eco* RII. A geometric decrease in amounts of higher order multimers from the monomer length is taken as a reflection of a random process of nucleotide divergence in regularly spaced restriction sites. Fry and Salser estimate that the same rate of divergence would produce a base change of 40–50% in the HS- α satellite. However, estimates of this kind are currently under suspicion for not only can mouse satellite be completely digested to monomer length under certain conditions with *Eco* RII and *Ava* II but also the different distribution of sites to other restriction enzymes suggests an apparent rate of divergence of $\sim 30\%$ in other portions of the satellite that also contain *Eco* RII sites (Horz & Zachau *Eur. J. Biochem.* **73**, 383; 1977).

In the light of such intriguing anomalies and in our ignorance of the full range of the behavioural quirks of replicating DNA we should tread cautiously before we reject arbitrary mechanisms of preservation and change despite the seeming difficulty in accommodating all the data.

For example, similarities in amplified sequences between species may result either from the presence of the same 'library' of starting sequences in the species group before amplification or because the species inherit a previously amplified sequence from a common ancestor. A comparison of the distribution of ten satellites in five of the seven sibling species of the *melanogaster* species subgroup with the phylogeny of the species based on the distribution of chromosomal inversions tells us that both processes may be operating (Barnes *et al. Chromosoma*, in the press). Another study in our laboratory on the distribution of satellites in species and subspecies (races) of the genus *Glossina* (tsetse fly) tells us that despite the proliferation of satellites in the group there are no obvious homologies between them and that neither process is operating. Instead, amplification might occur, subsequent to divergence, of sequences that are unique to each species. The data from Gramineae clearly suggests that amplification and maintenance of homogeneity continues, even after species divergence.

Differences in abundance of a satellite sequence common to several species may be the result of differences in the length of time during which the process of amplification of a member of a 'library' has taken place. Alternatively, differences in the chromosomal location of a sequence may be important in determining the rate at which amplification proceeds—a situation that has some support from the measured differential rates of sister chromatid exchange in different parts of the genome.

Similarities in minor sequence-variants in rodents are of course hardest to explain. They might be explained in two ways. For any given short repeated sequence there might be a restriction on types of mutational change that could occur (a situation that has some validity in the preferential transition of A $\rightarrow$ G in other DNA sequences) and so limit the numbers of sequence permutations in a six-base sequence to a few variants. Alternatively, it might be that there are only specific sequences to which the amplification process addresses itself—not in absolute terms but in terms of the positions of bases in a sequence in relation to other levels of chromatin organisation. The relic higher-order periodicities of the same length common to satellite DNAs of diverse mammalian species (Maio *et al. J. molec. Biol.* **117**, 637; 1977) are considered to be the result of nucleosome size determining the size of the presumed unit of amplification. In the case of the α satellite of African green monkey similarities in size are exact for some restriction endonuclease sites (at which illegitimate recombination might occur) fall at the interstices of adjacent nucleosomes (Musich *et al. J. molec. Biol.* **117**, 657; 1977).

Present-day distribution of sequence variants and rates of apparent divergence of satellite sequences might be a reflection of the differential registers of repeating DNA sequences and repeating histone octamers.

Finally, it is not improbable that amplification and interspersion of new with old sequences as seen in Gramineae and *Plethodon* are part and parcel of the same random process.

A process of computer simulation in our laboratory has shown that unequal exchange at the boundary of two different adjacent blocks of homogeneous sequences can generate a pattern of interspersed sequences determined by the initial unequal overlap.

It is always refreshing to relate new findings to the overall biology of organisms and to the processes of adaptation and speciation, and in this respect Fry's and Salser's interpretation of their data is highly pertinent and heuristic. For it can be argued that even if the processes responsible for the current distribution of satellites operate randomly, the arbitrary appearance of satellites and the ensuing change in karyotype could still conceivably be related to speciation. Bush *et al. (Proc. natn. Acad. Sci. U.S.A.* **74**, 3942; 1977) confirm a strong correlation between the rate of speciation and the rate of chromosome evolution in 225 genera of vertebrates. If it were to be shown that there is a direct causal relationship between the two, then we are left with the possibility, surprising as it may seem, that the initial steps in the genetic process of speciation are the result of accident rather than adaptation. □



A hundred years ago

A NOVEL and valuable application of electricity, designed to prevent the possibility of collisions on railways, is now the subject of experiment in the Marseilles station. It consists of an electric mirror, in which all the movements on a line 100 kilometres in length are brought vividly before the eye, and enables the station-masters to follow exactly the progress of every train. By this means it is hoped that all accidents resulting from delays or too rapid runs can be entirely avoided, and arrangements are being made for the general introduction into the stations of the new invention.

From *Nature* **17**, 7 March, 371; 1878.

Nuclear jet in a radio galaxy

from F. Graham Smith

QUASARS have a very potent source of energy, which may well be the gravitational field of a black hole. According to Lynden-Bell, the black hole is at the centre of a rotating galactic nucleus, so that matter falls in along two vortex lines at opposite ends of the spin axis. Outbursts of energy and radiation are emitted along the two axes, giving the linear patterns and central symmetry commonly seen in the radio emission from quasars and radio galaxies. We shall never see the black hole itself, but we can test the model by looking at the outbursts from it. A new test of this kind is to be found in recent observations of the elliptical galaxy NGC

6251, which are remarkable in the link which they provide between the very large scale structure of the radio galaxy and the very small scale structure of the nucleus itself.

A new aperture synthesis telescope at Cambridge, which works at the unusually long wavelength of 2 m, gave a map of a very large radio source, with angular extent of over 1° , centred on the galaxy (Waggett *et al. Mon. Not. Roy. astr. Soc.*, **181**, 465; 1977). The map had a resolution of 3.7 arc min, and it showed the typical double structure of a radio galaxy. The visible galaxy, only about 3 arc min across, coincided with one of the two radio lobes.

Optical measurements of the redshift, made with the Isaac Newton Telescope, gave the distance as 142 megaparsec (Mpc), so that the extent of the radio galaxy is about 3 Mpc. The next stage was to look at the galaxy with the better resolution of the Cambridge telescopes working at 21 cm, 11 cm and 2 cm. As the wavelength shortens the large radio galaxy fades, but extending from the centre of the elliptical galaxy there appears a remarkable jet of radio emission, very straight and narrow, about 200 kiloparsec (kpc) long, and pointing along the axis of one of the large radio lobes. At the shortest wavelength this jet could not be seen, and only an unresolved point-like radio source remained at the nucleus of the galaxy. This was less than 0.4 arcs across, that is, less than 16 kpc across.

Angular resolution down to 10^{-3} arc s is now available from Very Long Baseline Interferometry (VLBI). Observations by Readhead *et al.*, using radio telescopes at opposite sides of the US, have now resolved the point-like source, with astonishing results. This nuclear source has two parts, and again one is point-like and probably unresolved, while the other is again a narrow bright jet, pointing along the same axis.

These three successive mappings of the radio source have taken it apart as one might pull apart a pair of leaves in a plant bud, to find a smaller pair inside, and a further pair inside that. But the relative scale is quite different: the whole radio source is 3 Mpc across, the big jet is 200 kpc long, the nuclear jet is 1.7 pc across, while the widths of the nuclear jet and the nucleus itself are less than 0.1 pc. The conditions of radio emission are very different in the different sizes of object: for example the magnetic field in the nucleus must be of the order of 0.1 gauss (10^{-5} T), while in the outer radio lobes it must be about

10^{-6} gauss. Nevertheless there can be no doubt that the different regions are connected, and that they derive from the same energetic machine in the nucleus, whatever it may be. Attention is therefore given in these papers to the dynamics of the energy flow along the jets.

One additional feature is that the jets are seen only on one side. The explanation, again following Lynden-Bell, may be that the radio sources are moving with high relativistic velocities and that the axis along which they move is tilted towards our line of sight. We then only see the jet which comes towards us, the other being dimmed by a relativistic effect.

All this activity is to be found in an elliptical galaxy, 14th magnitude redshift 0.02, of fairly ordinary appearance optically. $\square$

Magnetic storm

from John Pendry

ONE of the basic properties of materials that solid state physicists set out to explain is the phenomenon of ferromagnetism. Because of the exclusion principle, electrons with parallel spins tend to stay away from one another and therefore have lower energy than anti-parallel spin electrons. This natural tendency for spins to align is opposed by the cost in kinetic energy of constructing parallel spin states, and by correlation effects which keep even anti-parallel spin electrons away from one another and reduce the advantage of alignment. The ferromagnetic state is a result of an interplay of these forces.

Despite this loose understanding of the forces at work great controversy still exists about what constitutes a correct description of a ferromagnet. At the centre of this controversy is a series of experiments in which electrons are ejected from nickel, a ferromagnet, and their spin polarisation measured. A strong magnetic field aligns the domains so these experiments tell us about the percentage alignments of the individual spins of electrons near the Fermi energy, which the ejection process selects. There are three ways in which electrons can be ejected: by photoemission (Bänninger *et al. Phys. Rev. Lett.* **25**, 585; 1970) by field emission (Landholt & Compagna *Proc. Int. Conf. on Physics of Transition Metals*, Toronto, 1977) and by electron capture to a fast ion whose trajectory grazes the surface (Eichner *et al. Proc. Int. Conf. on Physics of Transition Metals*, Toronto, 1977). All three experiments give wildly different spin polarisations and

until recently all three were in disagreement with theory. However high resolution experiments made in ETH Zurich (Eib & Alvarado *Phys. Rev. Lett.* **37**, 444; 1976) show a reversal of spin polarisation for electrons with energies very near the Fermi energy — a curious effect predicted on theoretical grounds by Wohlfarth (*Phys. Lett.* **36A**, 131; 1971) and a result which has intensified the theoretical debate.

One model of metals that has had great success in non-ferromagnetic metals is the single particle model in which the electrons move as independent particles—corrections for their interaction, for exchange and correlation appearing as an average field experienced by the electrons. This model has also been applied to ferromagnetic materials and explains the magnetisation in very simple terms: the mean field for the majority spin electrons has a strong exchange component and is therefore more attractive than that for minority spins. Therefore we have a self-consistent solution in which the majority create a deeper potential well for themselves which contains more electrons than the minority well. Because of its simplicity the model has the great virtue that it can make quantitative predictions in a parameter-free way. It can calculate amongst other things the magnetic moment per atom and the exchange splitting, that is, the increased depth of the majority potential well, both of which are accessible through experiment. The most advanced calculation of this nature has been published by Wang and Callaway (*Phys. Rev.* **15**, 298; 1977).

In opposition to the single particle theory strong arguments have been raised. It must be the case that electrons do not simply move around in a mean field. They can interact with fluctuations in the medium. For example an electron can reverse its spin and create a magnon (*Phil. Mag.* **24**, 203; 1971), a periodic disturbance in the direction of magnetisation. These virtual excitations if present in sufficient quantities will affect the experimental measurements of spin polarisation, and many believe that they do have a dominant role. The debate has not been resolved because the inclusion of these more complex effects is a formidable mathematical problem and no first principle predictions are possible in this realm. In fact, even for the single particle model, no rigorous connection between the theory and the three types of polarisation experiment has been published.

J. B. Pendry is Head of the Theoretical Physics Group at the Daresbury Laboratory.

F. Graham Smith is the Director of the Royal Greenwich Observatory, Herstmonceux.

The importance of making this connection is great because it will provide, through experiment, a severe test of the single particle model. If the model fails, we remain in the semi-quantitative era. If it succeeds the validity of calculations such as those of Wang and Callaway is established and the way opened for a detailed theoretical understanding of ferromagnetism. □

Stages in carcinogenesis

from Rod Langman

The Armand Hammer Cancer Workshop on Carcinogenesis was held at The Salk Institute, San Diego, from January 29–February 1, 1978

THE meeting started on a lively note with R. Peto (University of Oxford) presenting an epidemiological analysis of cancer in Britain.

We are well aware of the damning evidence linking smoking to lung cancer, and the latest data suggests a strong correlation between dietary fat (both animal and vegetable) and colorectal cancer. Preliminary tests have shown that these fats have no detectable mutagenic activity in the Ames bacterial assay.

In the epidemiological analysis of carcinomas, tumours of epithelial origin (which account for 90% of all cancers), there is a strong age dependence, with plots of log incidence against log age giving a straight line of slope 4–6 which can be interpreted as indicating 4–6 stages in the progression from normal to cancer. Peto emphasised that this relationship applied only to the carcinomas, and not to sarcomas where similar plots would tend to fall on a line with slope 1–2, but with considerable variation depending on particular tumour types. An interesting comparison was drawn between mice and men with the observation that both stand a roughly similar probability of getting cancer during their lifetime. Assuming man has 10^3 more cells at risk and lives 30 times longer, with the risk rising as the 4th power of age, $30^4 \times 10^3 = 10^9$ is the difference in probability of a single cell becoming cancerous in man compared with the mouse. Even if one assumes the number of potential targets to be the same in man and mouse, the age factor alone suggests a 10^{-6} differential.

Turning to studies with cultured

fibroblasts from man, mouse and hamster, the mutation frequencies at single loci are within a factor of 10^2 for both spontaneous and induced mutation. Thus, there seems to be no simple explanation for the difference between man and mouse in terms of mutation rate. Nevertheless, primary hamster fibroblasts after mutagenesis can form continuous cell lines, and show anchorage-independent growth in agar, in which case they also give tumours in animals (P. Ts'o, Johns Hopkins University). However, under essentially similar conditions of mutagenesis primary human fibroblasts do not give continuous cell lines that grow in agar and give tumours in nude (athymic) mice (R. Demars, Wisconsin). Perhaps mutagenesis *per se* is insufficient to cause all the changes necessary for a cell to become cancerous, and in the hamster example the culture conditions fortuitously supply those non-mutational factors in a specific manner.

The experimental phenomenon of initiation and promotion was analysed by B. Weinstein (Columbia University) who, in summarising the work in this field, pointed out several key findings. First, the initiator is a compound that usually behaves as a mutagen and can be given under conditions where no tumours form. Second, that promoters, which are not mutagens, and do not by themselves cause tumours, can cause the expression of tumours in animals pretreated with initiator. Among the various compounds that act as initiators urethane is interesting because it does not induce tumours at any dose tested, but does prepare an animal for the effects of promoters. Also it is possible to give initiator during the third trimester of pregnancy and observe the effects of promoter on the young adult progeny.

In contrast to the one-hit effect of initiators, promoters must be continuously applied to first cause papillomas, which disappear when promoter is removed, and only after prolonged application of promoter do tumours appear. In short, initiation and promotion provide what seems to be a good example of a multistage transition from a normal to a cancerous state. As Peto pointed out, the promoter-like effect of tobacco condensate is consistent with the epidemiological data on lung cancer in people who stop smoking. Within five years there is no further increase in lung cancer incidence. Furthermore, the effects of two agents such as smoking and asbestos on lung cancer incidence are not simply additive but multiplicative, again indicating different and independent sites of action in a multistage progression from normal to cancer.

Using a different approach L.

Siminovitch (Toronto), also emphasised a two-step process in the conversion of normal fibroblasts to tumours. Preparations of purified metaphase chromosomes from Chinese hamster ovary tumour cells could transform senescing (S^+) lung fibroblasts into immortal (S^-) cells, and although these S^- cells were also morphologically transformed (C^-), they would not give tumours in normal animals, and they maintained an anchorage-dependent (A^+) growth pattern. In a second step, using the same preparation of tumour chromosomes, these S^-C^- cells could be converted into tumorigenic cells which could grow as colonies in agar indicating a loss of anchorage-dependent (A^-) growth. At first glance, it would seem that conversion of normal to cancer cells requires two sequential steps, first S^+C^+ to S^-C^- , and second $S^-C^-A^+$ to $S^-C^-A^-$. However, one should be cautious of this kind of simplification because there are cases of S^+A^- cells in mouse (P. Patek, Salk) and human (R. Demars, University of Wisconsin).

One of the most interesting changes accompanying chromosome fragmentation-induced tumorigenesis was the marked decrease in chromosome stability as cells change from the $S^-C^-A^+$ to $S^-C^-A^-$ state; and this was not a general effect of chromosome transfer. Perhaps chromosome instability is a more general feature of highly tumorigenic cells giving them the potential to undergo rapid rearrangements in gene expression and thereby avoid many of the normal regulatory mechanisms.

No discussion of carcinogenesis can avoid the fascinating murine teratoma as a potential example of reversible carcinogenesis. K. Illmensee (University of Geneva) provided two new experiments. The first started with a single teratoma cell which was allowed to divide once in culture; one daughter cell was implanted in a blastocyst, while the other was tested for tumorigenicity in an adult mouse. Only in those cases where a viable mouse arose from the blastocyst and a tumour developed in the adult mouse were the experiments taken to be informative. There were 71 such cases; 18 of these were chimaeras, and five developed tumours sometime after birth; no chimaera developed a tumour. These findings raise several possibilities since it is well known that teratoma stem cells can undergo partial differentiation during tumour growth. For example, E. Reich (Rockefeller) argued that if one started with a stem cell, and upon division one daughter became committed to differentiation, only if that cell were chosen for implantation into the blastocyst would a chimaera develop and the other daughter cell give

R. Langman is Assistant Professor in the Laboratory of Developmental Biology at the Salk Institute.

a tumour. The converse case would not lead to a tumour growing in the adult mouse, and so would not be scored in a successful experiment. Of course, it is impossible to keep a complete balance sheet when the efficiency of tumour takes from a single cell is low. However, this kind of interpretation does not avoid the point that the teratoma can revert to normal at the single cell level, but that at the population level reversion to normal may be rare.

In another experiment provided by Illmensee a new field for the study of regulation in human cells was opened. He showed that a fusion between a mouse teratoma and human fibrosarcoma when implanted in a mouse blastocyst resulted in viable chimaeras which allowed the expression of at least one human enzyme—galactokinase. □

Ultracold neutrons

from M. Asghar

ULTRACOLD neutrons are slow neutrons with velocities of less than 7 m s^{-1} . In order both to study their properties and their interactions with matter and to use them for experiments in fundamental physics they must be stored and up to now this has raised various problems.

Charged particles such as protons and electrons can be contained in storage rings using electric and magnetic fields that interact with the electric charge carried by the particles. Although neutrons do not carry any net electric charge they can be stored in 'bottles' formed either of a magnetic field, in which the neutrons are contained by the interaction of their magnetic moment with the magnetic field, or (for ultracold neutrons) by materials which are very poor neutron absorbers, making use of the neutron-nuclear interaction.

Storage of ultracold neutrons in bottles made of matter is possible because if the neutron wavelength is much greater than the interatomic distance of the containing matter the average (effective) potential due to the neutron-nuclear interaction is $\sim 10^{-7} \text{ eV}$ and is repulsive for many materials. Neutrons with energies $\lesssim 10^{-7} \text{ eV}$ (velocity $\lesssim 7 \text{ m s}^{-1}$) are therefore totally reflected from the bottle walls and can be contained.

The ultracold neutrons form a small fraction ($\sim 10^{-11}$) of the thermal neutron spectrum in a reactor. Totally reflecting curved guides installed either horizontally, vertically or in an inclined position with respect to the reactor core, have been used to bring out ultracold neutrons from reactors whilst removing the fast neutrons

and accompanying gamma rays. The ultracold neutrons are guided out from the reactor by total reflection from the specially prepared walls of the guide just like photons in a light guide (Luschikov, *Proc. Int. Conf. Interaction of Neutrons with Nuclei*, Lowell, Massachusetts, 1976 1,117; *Physics Today*, 42, June 1977). Ultracold neutrons can also be obtained from relatively cold neutrons ($V \sim 50 \text{ m s}^{-1}$) by slowing them down mechanically by successive reflections from a system of rotating mirrors called a neutron turbine (Steyerl, *Nucl. Inst. Meth.*, 125, 461; 1975). The energy of these neutrons can be easily selected by using the Earth's gravitational field; for example, 10^{-7} eV neutrons will rise only $\sim 1 \text{ m}$ in height.

More than a dozen laboratories in the world are working on and with ultracold neutrons. Most are in the USSR, where the work on the ultracold neutron storage was originally started. The ultracold neutron flux ϕ_{UCN} available at these laboratories is $\sim 1\text{--}10 \text{ n cm}^{-2} \text{ s}$. However, the recently installed UCN facility at the Grenoble high flux reactor has $\phi_{\text{UCN}} \sim 250 \text{ n cm}^{-2} \text{ s}$, with an ordinary water source. This ϕ_{UCN} can be increased by a factor of ~ 20 by using helium cooled sources (Smith, *ILL-Grenoble Workshop*, October 1977).

Bottles made of different low absorption materials such as copper, graphite, quartz, teflon, beryllium, glass and stainless steel have been used to store ultracold neutrons. However, the loss of neutrons due to processes other than the decay of a free neutron ($T_{1/2} = 10.61 \text{ min}$) was found to be always much higher (by a factor of ~ 100) than the theoretical estimates. Moreover, the experimental data showed that the storage time (a few hundred seconds maximum) was independent of the container material. Detailed experimental work led to the belief that this loss must result from the interaction of neutrons with the walls of the bottle. It could be due to the presence of an impurity with a high neutron absorption cross section such as ^{10}B ; but surfaces cleaned and polished in different ways made little difference to the loss rate of neutrons. Another possibility is the presence of a hydrogen surface layer, for example, which could inelastically scatter a neutron and increase its energy above the effective potential step of the wall material. That neutron would not remain stored. If an impurity such as hydrogen was the cause, changing the temperature of the bottle should influence the neutron storage time. However, extensive work on bottles of different materials in a temperature range of about -170 to 500°C did not modify the neutron loss rates significantly (Groshev *et al. Proc. Conf. Neutrons*, Kiev, 1975). The reluctant conclusion was that the anomalously high loss of neutrons could not be due to the presence of surface impurities.

Recently Lanford and Golub (*Phys. Rev. Lett.* 39, 1509; 1977) have measured the hydrogen concentration profiles on copper, graphite and glass surfaces. The different surfaces were prepared and used in a way similar to that used for the bottles. The hydrogen density profile measurements were done by measuring the 4.43 MeV γ -ray yield from the $^{15}\text{N} + \text{H} \rightarrow ^{12}\text{C} + ^4\text{He} + 4.43 \text{ MeV}$ γ -ray resonant reaction. The hydrogen concentration as a function of depth in the sample wall up to $\sim 100\text{--}170 \text{ \AA}$ (comparable to the ultracold neutron penetration in these solids) was determined by varying the energy of the incident ^{15}N ions. All these samples were studied both at room temperature and after heating for various times at different temperatures ($190\text{--}265^\circ \text{C}$). These results show that the hydrogen surface concentration of these samples is about the same: $\sim 10^{23} \text{ H atoms cm}^{-2}$, within a factor of about three and it decreases by a factor of only about two, when they are heated. The average probability of loss of an ultracold neutron per reflection deduced from these data is close to the existing experimental values for these surfaces. These results indicate strongly that the anomalously high loss of ultracold neutrons in these bottles is quite consistent with a high concentration of surface hydrogen an interpretation also suggested by mirror reflection of ultracold neutrons (Schreckenhofer & Steyerl, *Phys. Rev. Lett.*, 39, 1310; 1970). To be absolutely certain of this conclusion a hydrogen free surface must be produced and the effect on the storage time measured. That will not be easy. Any decrease in hydrogen surface density due to heating will tend to be compensated by a higher inelastic scattering cross section. Hydrogen may be so strongly bound in the surface that the vacuum and the temperatures used are not enough to dislocate it or there might be enough hydrogen in the material that any depletion at the surface is compensated by diffusion of the in-depth hydrogen. Hence to obtain long storage times, contamination of the bottle walls will have to be reduced and this will need a detailed study of surfaces. One possibility is to clean surfaces by electron bombarding in a clean and good vacuum ($\sim 10^{-10}\text{--}10^{-11} \text{ Torr}$) and depositing a layer of another material such as Ni. Alternatively replacing the surface with a very pure liquid ^4He , or a solid oxygen layer can be tried: these materials have a low neutron absorption. This operation probably will have to be done *in situ* and will form a part of the storing bottle.

Storage of ultracold neutrons is required so that the properties of the neutrons themselves can be studied as well as their interaction with matter. A relatively good flux of ultracold neutrons could provide monochromatic beams for scattering experiments. But in addition, a

M. Asghar is a physicist at the Institute Laue-Langevin, Grenoble, France.

number of fundamental physics experiments are possible.

The free neutron decay half-life

The neutron decay half-life is an important quantity in the weak polar vector and axial vector coupling constants characterising β -decay. This parameter is also important in astrophysics and in calculating the structure of neutron stars. The currently accepted value of 10.61 ± 0.16 min (Christensen *et al.*, *Phys. Rev. D* **5**, 1628; 1972) has only a precision of $\sim 1.5\%$. It would be worthwhile to improve this value by as great a factor as possible. The ultracold neutrons stored in a bottle may be one way; but this method will be practical only if one can reduce the neutron losses by a large factor (~ 50 – 100 , say).

The neutron electric dipole moment (EDM)

The existence and observation of the neutron EDM will violate the symmetry laws of parity (P) and time (T) simultaneously. The neutron EDM has been sought for the past 30 years using principally the neutron beam magnetic resonance spectrometer. In this method polarised neutrons are subjected to uniform electric and magnetic fields in the presence of suitably oscillating magnetic fields. The present upper limit of EDM is $\mu_n < 3 \times 10^{-24}$ e cm where e is the electron charge (Dress *et al.*, *Phys. Rev. D* **15**, 9; 1977). This experiment was done at the Grenoble high flux reactor. This value rules out the calculated EDM values based on the weak and the electromagnetic interactions. The bulk of the multiweak theories lie just above the present limit, whereas the higher order milliweak, millistrong and superweak theories predict EDM down to $\sim 10^{-30}$ e cm.

This EDM limit can be improved by using a neutron bottle. With a neutron containment time of 30 s and an electric field of 30 kV cm^{-1} , one reaches an EDM sensitivity of 2×10^{-26} e cm with 3 months of data taking. Such an experiment is being prepared at the Grenoble ultracold neutron source (Dress & Miller, *ILL Workshop*, October, 1977). If one can improve considerably the ultracold neutron storage time and their density (at present only $\sim 0.5 \text{ cm}^{-3}$) by, for example, using a bottle filled with liquid helium as recently suggested by Golub and Pendlebury (*Phys. Lett.*, **62A**, 338; 1977), it might be possible to reach a sensitivity of a few times of 10^{-28} e cm. Such an experiment is also being carried out in Leningrad by Lobashov *et al.*

Parity and time violation in the neutron's gravitational interaction

One knows that the weaker interactions are more symmetry violating than the stronger ones. Does the gravitation interaction itself violate any symmetry? For example, one can look for the

gravitational dipole moment (GDM) of a particle such as neutron. The existence of a non-zero GDM for neutrons will imply the relative displacement between the gravitational and inertial masses of a particle, that is, the breakdown of the equivalence principle, and again the violation of the parity and the time symmetries. The interaction potential due to the Earth's field is $V = 2 \times 10^{-23} \alpha (\vec{\sigma} \cdot \hat{r})$ (eV) (Golub, *ILL Workshop*, October, 1977), where α is proportional to the GDM of the neutron, and $\vec{\sigma}$ and $\hat{r}$ are respectively the spin of the neutron and the unit vector in the direction of the Earth's gravitational field. The present experimental limit on α is $\sim 10^8$. If one attempts an experiment similar to an EDM experiment with a sensitivity of 10^{-27} e cm, one should obtain a limit on $\alpha \sim 1$, if the magnetic field is reversed with respect to the Earth's gravitational field.

Ultracold neutron sources and bottles are still in their early stages of development. Many problems have to be tackled in order to obtain high neutron density and long storage times. However, their properties are beginning to be understood. $\square$

Protocols for computer networks

from Roland Rosner

THE idea that access to a computer system is restricted only to those who work in the same building or can be connected by a fixed link is gradually being buried by progress on computer networks. Similar to the telephone system, these provide a communications grid which enables all the attached terminals and computers to exchange information.

During the past few weeks, two significant news items on computer networks have received undeservedly scant attention in all but the specialist press. The first was an announcement by the UK Post Office of its initial, admittedly tentative, steps towards the establishment of a national network geared specifically to the transmission of data. This was followed shortly afterwards by publication of the tariffs for the EEC's data network EURO-NET, which is to go into service next year as a means of giving community-wide access to scientific and commercial databases located in the

member countries.

The operation of the new public data networks now being proposed by the postal administrations will be based on the technique known as packet-switching. In contrast to the conventional circuit-switched telephone system, this requires the fragmentation of data for a given call into short 'packets' which share the bandwidth of the network's transmission lines with packets from other unrelated calls. Within the network, the exchanges must receive, examine and forward the packets along routes determined by the addresses contained in the packets themselves. The subscribers' equipment must possess sufficient intelligence to perform the fragmentation and reassembly operations at both ends of a call. The rules for connecting equipment to such networks are enshrined in an international standard which was agreed by the world's postal administrations in 1976 with an unprecedented sense of urgency.

However, adherence to the attachment rules is sufficient only to establish a liaison or call between a pair of subscriber computers or terminals and to guarantee a high probability that packets will be delivered in sequence and free of transmission errors. It does not go as far as providing for the meaningful exchange of information within the call. By analogy with the need for a common language if the meaning of a letter is to be communicated from sender to recipient, common procedures must be agreed upon beforehand and executed by both participants during the call so that each has the same view of the structure underlying the data flowing between them. Such procedures are known, appropriately enough, as protocols and their importance may be gauged from the fact that a major international conference, held recently in Liege, Belgium, was devoted entirely to this topic.

In his introduction to the conference proceedings, Christopher Layton from the Commission of the European Communities points out that, until recently, computers have been run under the supervision of highly autocratic operating systems and communicate with a fixed population of dedicated terminals by means of proprietary protocols. These restrict evolution and ensure incompatibility among the equipment from different manufacturers. Real benefits will only be derived from the new networks if computers and terminals are able to communicate with each other regardless of pedigree. This will require the specification, standardisation and implementation of the appropriate protocols.

An example of the increasing attention being given to the influence of

Roland Rosner is a Principal Scientific Officer at the Rutherford Laboratory and a member of the Network Unit of the Computer Board and Research Councils.

networks on computer systems is a paper presented by F. Hertweck of the Institute of Plasma Physics at Garching, Germany. It describes a practical method for incorporating the standard call and data communications procedures within a modern operating system and for the sharing of a single physical link by traffic to different destinations from several processes running simultaneously in one computer. The approach adopted is to structure the software into a number of independent layers each of which uses facilities provided by lower layers without concerning itself with the mechanisms they employ. This architecture has gained widespread acceptance among the designers of the advanced protocols now under consideration. One of these is a file transfer protocol prepared by a UK group working under the aegis of the Post Office's experimental network project EPSS. Completed too late for presentation in Liege, the protocol handles the shipping of large volumes of data across a network and is flexible enough to deal with serious mismatches in the processing, communications and input/output facilities at either end of the call. The protocol is to be implemented initially on computers in UK universities and research establishments and, after a period of intensive testing, it will be a candidate for adoption as a standard.

The vast variety of visual display units on the market makes the problem of devising protocols for communication between computers and such terminals particularly complex. No real contributions towards a solution were made at Liege though this is, in many ways, among the most urgent of requirements. The imminent opening of EURONET, with its initial emphasis on large-scale terminal access to databases, carries with it the danger that a hastily conceived protocol will, by virtue of its widespread adoption, become an unsatisfactory *de facto* standard.

The conference indicated how much work still remains to be done if the paths of the postal administrations, the users and the manufacturers of computers and terminals are to converge. By definition, studies on protocols for computer communication cannot be conducted in isolated ivory towers but must be coordinated at all levels. Liege confirmed that UK expertise is among the world leaders and it would be prudent to capitalise on this by formulating a centrally funded national plan for protocol development. □

Nuclear matter distributions

from P. E. Hodgson

THE spatial distribution of protons and neutrons in the nucleus is an important feature of nuclear structure, and also has a strong influence on the rate of nuclear reactions. The distribution of charge, and hence that of the protons is known rather accurately for many nuclei from analyses of electron elastic scattering, but it is only recently that it has proved possible to measure the neutron distribution with anything approaching the same precision.

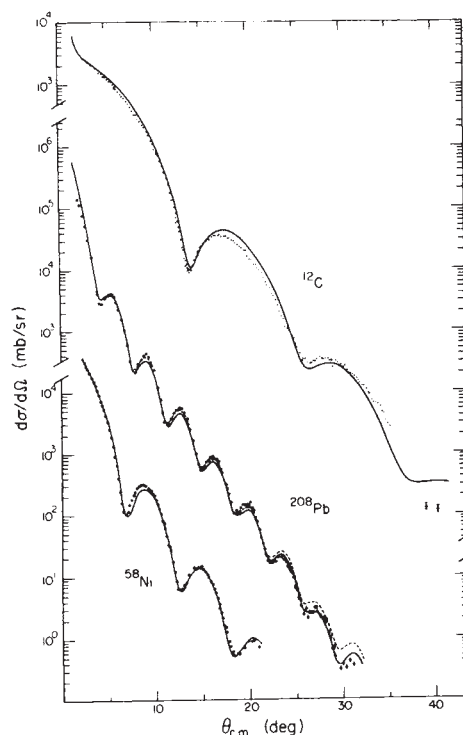
The elastic scattering of high-energy electrons from nuclei can easily be analysed because the electromagnetic interaction is well understood. The calculations are quite complicated but can easily be handled by fast computers. The method used is to postulate a nuclear charge distribution of a reasonable shape, usually a three-parameter Fermi distribution, and then to adjust the parameters to optimise the agreement between the measured electron elastic scattering cross section and that calculated from the postulated charge distribution. Very sophisticated methods of doing this have been developed during the past few years, and now the charge distributions of many nuclei are known with impressive accuracy.

The neutron distribution cannot be found in the same way because the nuclear force is not so well understood as the electromagnetic force. The elastic scattering of protons and neutrons by nuclei can be measured, but the calculation of the scattering to be expected from an assumed neutron and proton distribution is a very difficult problem that has still not yet been solved with sufficient accuracy for the present purpose.

A way round this difficulty that has been developed recently is to measure the elastic scattering of very high energy protons from the nucleus whose nucleon distribution is required. If the energy is high enough, and this means about 1,000 MeV, the scattering can be calculated by treating the nucleus as an assembly of individual nucleons. This is justified because at such energies the wavelength of the incident protons is small compared with the distance between the nucleons in the nucleus. At lower energies, where this is not so, the scattering takes place from the nucleus as a whole, and this gives rise to the computational difficulties already mentioned.

Some new work on these lines has

P. E. Hodgson is a Lecturer in Nuclear Physics at the University of Oxford.



Differential cross sections for the elastic scattering of 800 MeV protons by ^{12}C , ^{58}Ni , and ^{208}Pb compared with theoretical calculations.

recently been completed by Blanpied and colleagues (*Phys. Rev. Lett.*, **39**, 1447; 1977). They measured to very high accuracy the cross sections for the elastic scattering of 800 MeV protons from ^{12}C , ^{58}Ni , and ^{208}Pb using the high-resolution spectrometer at the Clinton P. Anderson Meson Physics Facility at Los Alamos. The results are shown in the figure, compared with theoretical curves calculated using the proton-nucleus optical potential obtained from the formalism of Kerman, McManus and Thaler, with a general spin-dependent form for the nucleon-nucleon amplitudes. The parameters of the expressions for these amplitudes were determined by fitting experimental results for proton-nucleon scattering. The nucleon distribution was chosen to have a three-parameter Fermi form with a Gaussian correction term. The formalism then allows the proton scattering cross section to be calculated from any assumed nucleon distribution, and the analysis can then proceed as in the case of electron scattering. The comparison in the figure shows that the experimental data can be fitted very accurately, and this suggests that the nucleon distribution is quite well determined.

The paper gives the parameters of the optimum nucleon distributions. These show that the difference between the root mean square radii of the proton and neutron distributions in the nucleus is not significantly different from zero, which is in line with a variety of other estimates of these

quantities. Some more detailed calculations using the Gaussian correction term in the nucleon distribution showed how sensitive the analysis is to small details of the distribution. The solid curve in the figure for ^{208}Pb shows the improvement that can be obtained by including this term, compared with the dashed curve when it is omitted.

These calculations are only preliminary and will certainly soon be improved. In particular, the use of parametrised nucleon distributions is known to introduce undesirable correlations between the densities at different radial distances, and it is preferable to use either an expression with no such correlations, or a charge distribution with a more fundamental theoretical basis. Work along these lines is in progress.

Similar analyses of the cross sections for the elastic scattering of 1,000 MeV protons measured by the Saclay group are being made by Varma and Zamick at Rutgers (*Phys. Rev.*, **C16**, 308; 1977) and it will be of great interest to compare the results of the two groups. All this work will certainly lead to more precise knowledge of the distribution of nucleons in the nucleus. □

Dilatancy has magnetic effect

from Peter J. Smith

WHEN a stress is applied to, or removed from, a rock there is a distortion of crystal structure which often gives rise to a small change in the rock's remanent magnetisation. This is known as the piezomagnetic effect and has an obvious application to the prediction of earthquakes. The variations in upper crustal stress which precede a seismic event should manifest themselves as changes in crustal magnetism and hence in that small part of the local geomagnetic field arising from permanent rock magnetisations. Continuous monitoring of the geomagnetic field should therefore lead to the observation of geomagnetic variations which reflect subsurface stress changes occurring before the onset of an earthquake.

At least, that is the theory. In practice, piezomagnetism has been of little use in earthquake prediction, partly because very small field changes are difficult to measure, partly because piezomagnetic changes tend to be swamped by normal geomagnetic noise, and partly because magnetometers are liable to be jolted by seismic vibrations and thus indicate spurious field changes. Most of the seismic-geomagnetic correlations reported since 1799 are there-

fore certainly not valid. Moreover, even when the practical problems have been solved by carefully designed investigations with modern techniques and equipment, the piezomagnetic effect has often proved elusive in the field. Breiner (thesis, Stanford University, 1967), for example, was unable to discover any seismomagnetic changes preceding Californian earthquakes, although he did manage to correlate local magnetic changes with creep events on the San Andreas fault during the period 1965-67.

Following the publication of Breiner's disappointing results, interest in seismomagnetism withered as attention shifted to newly-discovered and apparently more promising phenomena such as premonitory variations in the seismic velocity ratio. But in recent years there has been a minor revival, with possible seismomagnetic effects being reported by Johnston *et al.* (*Seismol. Soc. Am. Bull.*, **65**, 1129; 1975) and B. E. Smith and Johnston (*J. geophys. Res.*, **81**, 3556; 1976) in connection with Californian earthquakes. At the same time there has been a change in the theoretical background against which the piezomagnetic effect in Earth materials must be viewed. For as Revol *et al.* (*Earth planet. Sci. Lett.*, **37**, 296; 1977) point out, most laboratory studies have been concerned with stresses in the elastic range, a reflection of the view that earthquakes are the result of elastic processes. The more recent dilatancy model offers a new perspective, however. Before failure, rocks undergo an inelastic volume increase which is believed to result from the formation and propagation of cracks; and although the volume of rock subject to near-breaking stress before an earthquake is likely to be small, it may nevertheless 'hold the key to prediction because the approach to failure may be accompanied by detectable time-dependent magnetic effects'.

With this in mind, Revol and his colleagues have carried out an exploratory series of experiments in which small cylindrical rock samples submitted to isothermal uniaxial compression without confining pressure were gradually stressed to failure. At the same time both induced and remanent magnetisations in the rocks were monitored continuously and their variations recorded. The results were very varied, depending on the rock type (natural magnetite, serpentinite, granodiorite) and on the particular magnetic parameter (susceptibility, saturation isothermal remanent magnetisation, thermoremanent magnetisation, and so on). Nevertheless, it is evident that dilatancy does have a magnetic expression in many cases.

As far as the natural magnetite is

concerned, the susceptibility parallel to compression decreases with increasing stress, but the onset of dilatancy (at about half the breaking stress) is marked by a reduction in the rate of decrease and just before failure there is a small anomalous increase in susceptibility. Increasing stress also generally results in a decrease in remanent and induced magnetisations, again with noticeable changes at the onset of dilatancy. Thermoremanent and partially demagnetised isothermal remanent magnetisations are unexpected exceptions, however, for these initially increase with increasing stress but then decrease in such a way that they return to their original values when dilatancy begins. For all types of magnetisation there are small but inconsistent magnetic changes immediately before failure.

Some serpentinite samples exhibit magnetic changes with stress and some do not, whereas none of the granodiorite samples does. The deciding factor here seems to be the domain state of the magnetic minerals, multi-domain material such as the natural magnetite reacts to compression whereas single domain material remains unaffected, although it is not clear whether the smaller single domain grains are incapable of reacting to stress or whether they simply do not see the stress before failure is reached. Where magnetic changes do occur, however, they are generally similar to those in magnetite. In other words, they always indicate in some way the onset of dilatancy and there are always small anomalous effects close to failure, enabling the failure to be predicted.

Clearly these experiments are only a beginning. The initial increases in thermoremanent and partially demagnetised isothermal remanent magnetisations obviously require further study. And more generally, Revol *et al.* plan experiments to stimulate the temperatures and confining pressures found at depth in the Earth's crust. But whether any magnetic effects seen in the laboratory will be detectable in the field is another matter. Past experience suggests that a pessimistic view must be taken, although Revol and his co-workers do offer one ray of hope. They find that the application of stress to a rock also gives rise to magnetic directional changes which are, if anything, more marked than the changes in total magnetisation. This suggests that three-component field observation may be rather more successful than total field intensity observation in detecting piezomagnetic effects in the Earth's upper crust. □

Peter J. Smith is a Reader in the Department of Earth Sciences at the Open University.

articles

A jet in the nucleus of NGC6251

A. C. S. Readhead, M. H. Cohen & R. D. Blandford

California Institute of Technology, Pasadena, California 91125

A core and a jet of length ~ 1.7 pc have been found in the nucleus of NGC6251. The nuclear jet is aligned with both the 200-kpc jet and the outer lobes, which have a separation of ~ 3 Mpc; the alignment persists over a range of scales $\gtrsim 5 \times 10^6:1$. The observations provide strong evidence that energy and momentum are transferred continuously from the central core to the outer lobes by a beam.

AN exceptionally long and narrow radio jet has recently been discovered in NGC6251 (ref. 1), a 14 mag elliptical galaxy with $z = 0.023$. The jet is aligned with two very large radio lobes and is perpendicular to the major axis of the galaxy. The nucleus of the galaxy coincides, to within the accuracy of the sky survey prints, with a compact radio source at one end of the jet. The nuclear source is unresolved by the Cambridge 5-km telescope at 15 GHz and therefore has an angular size < 0.4 arc s. In this article we report observations which show that the nuclear source itself contains a jet which is aligned with the outer jet and we discuss the physical conditions in the radio-emitting regions.

Observations

The observations were carried out on 23–24 July 1977, using telescopes at the Haystack Observatory (HSTK) in Westford, Massachusetts, the National Radio Astronomy Observatory (NRAO) in Green Bank, West Virginia, and the Owens Valley Radio Observatory (OVRO) in Big Pine, California. Several sources were observed, but here we report only on NGC6251.

The observations were made at $\lambda = 2.82$ cm using left circular polarisation. IF data were recorded with the standard Mk-II VLBI system, at 2-MHz bandwidth² and the video tapes were correlated on the CIT/JPL processor in Pasadena.

We obtained substantial improvement in sensitivity over previous observations at 2.8 cm, as all three stations had cryogenic amplifiers and hydrogen maser oscillators. Two of the amplifiers were new (at HSTK and OVRO) and these are the first results to be reported with them. Table 1 lists the characteristics of the systems.

The values for the system temperature, T_{sys} , and degrees of antenna temperature per Jansky are optimal values, and both get substantially worse at high zenith angles. We used a fringe-fitting program written by G. H. Purcell, and integrated coherently for 4 min. This gave a statistical noise of 30–40 mJy, but there are larger uncertainties in the correlated flux density due to scaling errors. The statistical errors in phase depend on the correlated flux density and range from 3° to 8° . The phase is independent of scaling errors.

The Mark II system uses 1-bit recording and corrections must be made for the ratio of antenna temperature to system temperature, and for various approximations used in the correlator. We generally followed the method outlined by Cohen

*et al.*³. The procedure uses 'bootstrap' adjustments to attain consistency, and we have only been able to do this partially because the full observations, on all sources, have not yet been reduced. Thus the amplitudes we report here are provisional, but we expect that any changes will be less than 10%.

The flux density scale is based on Virgo A, for which we assumed $S_{2.8} = 38.0$ Jy (ref. 4). NGC6251 itself was too weak for accurate measurement during the observations, and we assumed its total flux density was 0.9 Jy, the value reported by Waggett *et al.*¹. A complication is that the source is at high declination ($\delta = 82.6^\circ$). None of the telescopes had calibration data for that part of the sky, but it is unlikely that unknown gain and pointing errors amount to more than 5%.

In Fig. 1a and b we show the amplitudes of the visibility function on the three baselines, and the closure phases. A closure phase is the sum of the three visibility phases around the triangle at a common epoch, and is independent of oscillator and atmospheric instabilities^{5,6}. The error bars on the amplitudes are the quadrature combination of statistical noise and a 5% allowance for systematic effects (such as pointing error). In addition, there is a baseline-to-baseline scaling error which could be as big as 10%.

Model Fitting

In model fitting we used both the fringe amplitudes on the three baselines and the closure phases. The latter eliminate the 180° position angle ambiguity, which is present when amplitude data alone are used, and enable us to determine the correct orientation of the source.

In choosing the parameters of the model we were guided by the following considerations: the fringe amplitudes, γ , on the two longer baselines (NRAO–OVRO and HSTK–OVRO) show clearly that the source is well resolved in the position angle (pa) of the outer jet ($\sim 120^\circ$), and practically unresolved in the direction perpendicular to this (that is in pa $\sim 30^\circ$, interferometer-hour-angle ~ 4 h). The mean of the peak visibility on these two baselines is $\bar{\gamma} = 0.97$. Assuming a gaussian brightness distribution, this corresponds to a full width at half maximum of 0.13 milli arc sec (arc ms), and we therefore adopt this value in our final model (see Table 2). The closure phases (Fig. 1b) depart significantly from zero, and show that the nuclear source is asymmetric and has complex structure. We find that it consists of two components with roughly equal flux densities. It is therefore possible that one component has a maximum visibility very near unity in pa 30°

Table 1 System characteristics

	HSTK	NRAO	OVRO
T_{sys} (K)	77	120	77
T_{ant} (K Jy ⁻¹)	0.13	0.23	0.19

on the longer baselines, and the other has $\gamma_{\max} = 0.94$, that is $\text{FWHM} = 0.18$ arc ms. However, in view of the 10% uncertainty in the calibration, there is no firm evidence that the source is resolved in pa 30° , since the observed maximum visibility on each of the longer baselines is consistent with unity. It is likewise possible that the maximum visibility has been overestimated by up to 10%, in which case,

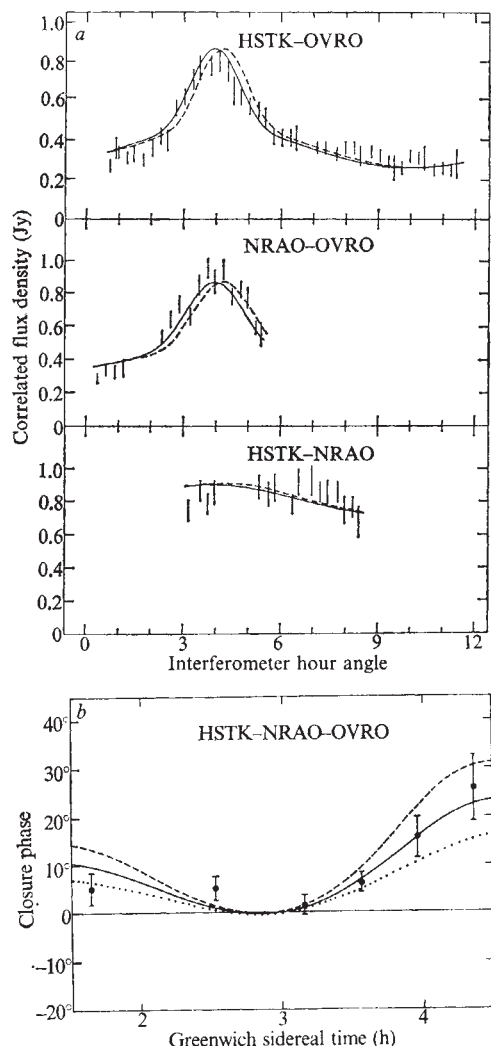


Fig. 1 *a*, The correlated flux densities observed on the three baselines. The solid line shows the fit of the adopted model (pa = 300.5° , see Table 2) to the data; the dashed line shows the fit for a model in the same position angle as the outer jet (pa = 296.5°). Data from 24 July 1977; $S_T = 0.9$ Jy. *b*, Comparison of the observed closure phases with the closure phases predicted by three models: (1) The adopted model (Table 2), separation between core and jet centres = 0.85 arc ms (solid line). (2) Model with identical parameters to (1) except that the core-jet separation is 0.60 arc ms (dotted line). (3) Same as (2), but with a core-jet separation of 1.1 arc ms (dashed line).

if only one component is resolved, we may place a firm upper limit of 0.5 arc ms on the width of this component.

On either side of pa 30° , on the cross-country baselines, the visibility drops very rapidly, indicating that one of the components is a highly elongated nuclear jet with pa in the range 119° – 123° . There is, however, a long tail to the visibility function (see Fig. 1*a*), indicating that the other component is a small core which is only moderately resolved in pa 120° . In this position angle $\gamma = 0.26$ on the longest baseline, thus $\gamma_{\text{core}} \sim 0.52$ in pa 120° , giving a width of 0.63 arc ms. This is the value adopted in the final model (see Table 2). If we again allow for a possible overestimate, by 10%, of the visibility amplitudes we derive a firm upper limit of 0.7 arc ms for the elongation of the core along the position angle of the nuclear jet.

The detailed model fitting to the visibility amplitudes and closure phases was done separately. We found that these data were, to some extent, complementary in the limits they imposed on the model—the closure phase data giving the strictest limits on the separation of the two components, and the visibility amplitudes doing the same for the dimensions of the core and jet. In fitting the amplitudes, we assumed that the nuclear source consists of two elliptical gaussian components. However, it is clear from the amplitude data that the nuclear source is basically one-dimensional, thus in model fitting to the closure phases we assumed a one-dimensional model consisting of two gaussian components.

In both fitting procedures we varied each of the model parameters and chose the optimum values by the method of least squares (Table 2). In each case the limits indicate the range of acceptable values for the parameter.

The final model adopted represents the best fit to both the amplitudes and the closure phases together (Table 2). The model accounts for all the flux density of the nuclear source. The fit of the adopted model to the data is shown by the solid lines in Fig. 1*a,b*. We are confident that it is basically correct, within the limits set by the restricted (u, v) coverage, and we do not think that it will be appreciably changed by any recalibration, when the remainder of the observed sources are processed.

In Fig. 2 we show the main features of the radio source covering a range $> 10^6:1$ in scale. The contour diagrams of the whole source and outer jet are taken from Waggett *et al.*¹, while that of the nuclear jet represents the adopted model. At a distance of 142 Mpc (ref. 1), the projected length of the nuclear jet (FWHM) is 1.7 pc, and the transverse dimension is 0.09 pc.

The position angle of the nuclear jet relative to the core, $300.5 \pm 2.0^\circ$ is remarkably close to the position angle of the outer jet ($296.5 \pm 0.5^\circ$) observed with the Cambridge 5-km telescope¹. As a comparison we also show, in Fig. 1*a*, the fit of the adopted model rotated through 4° to the same position angle as the outer jet (dashed line).

Physical conditions in radio-emitting regions

Characteristic dimensions, equipartition fields, and electron cooling times for the different radio-emitting regions are given in Table 3.

Nuclear core: the combined spectrum of the core and nuclear jet shows a turnover at ~ 3 GHz. If we attribute

Table 2 Parameters of adopted model

	Limits set by amplitude data	Limits set by closure phase data	Adopted model
Position angle of jet relative to core	$121 \pm 2^\circ$ or $301 \pm 2^\circ$	$299 \pm 3^\circ$	300.5°
Separation of core from jet centre (arc ms),	< 1.0	0.85 ± 0.25	0.85
Flux density of core	1.2 ± 0.5	1.4 ± 0.5	1.3
Flux density of jet			
Dimensions of core (arc ms)	$< 0.7 \times < 0.5$	< 1.0	0.63×0.13
Dimensions of jet (arc ms)	$2.2 \pm 0.5 \times < 0.5$	2.8 ± 0.5	2.5×0.13

Table 3 Conditions in the various radio-emitting regions of NGC6251

	Nuclear core	Nuclear jet	Outer jet	Hot spot*	Tail†
Distance from core centre (r kpc)	2×10^{-4}	2×10^{-3}	100	600	400
Diameter, assuming cylindrical symmetry (d kpc)	9×10^{-5}	9×10^{-5}	5	80	200
Brightness temperature at 5 GHz (T K)	5×10^{11}	8×10^{10}	20	0.4	0.09
Minimum energy density (u erg cm $^{-3}$)‡	2×10^{-3}	10^{-3}	6×10^{-12}	10^{-13}	4×10^{-14}
Equipartition magnetic field (G)	0.2	0.1	8×10^{-6}	10^{-6}	6×10^{-7}
Cooling time for 5 GHz-emitting electron in equipartition field (yr)	6	10	2×10^7	8×10^7 §	8×10^7 §

* Hot spot in western component at $\alpha = 16$ h 33 min, $\delta = 82^\circ 48'$.

† Low surface brightness tail around $\alpha = 16$ h 35 min, $\delta = 82^\circ 45'$.

‡ Calculated from the formula $u = 2.6 \times 10^{-12} (T/d)^{4/7}$ where we have assumed a spectral index of -0.5 extending over three decades of frequency and a negligible energy density associated with the proton and thermal components. A spectral index of 0 and an upper cut-off of 500 GHz are assumed for the core component.

§ Cooling time in the microwave background.

this to synchrotron self-absorption of the nuclear jet, then the present observation is compatible with a flat-spectrum core turning over at ~ 5 GHz. In addition, if we assume that the core is marginally resolved in both dimensions, its brightness temperature is $\sim 5 \times 10^{11}$ K, just below the 'inverse Compton' limit⁷, and the magnetic field strength is consistent with the equipartition value ~ 0.2 G (ref. 8).

Nuclear jet: if we adopt a spectral index ~ -0.5 (as in the outer jet¹), and a width ~ 0.13 arc ms, then the turnover at ~ 3 GHz can also be interpreted as synchrotron self-absorption with a magnetic field strength close to the equipartition value. The minimum energy density provides an approximate lower bound to the pressure ($\gtrsim 2 \times 10^{-3}$ dyn cm $^{-2}$) exerted by the jet on its surroundings and it is of interest to enquire what constraints are imposed on a hypothetical gas cloud of radius ≥ 1.5 pc in pressure equilibrium with the jet. This cloud must be optically thin to Thomson scattering and free-free absorption and not have a detectable X-ray flux at the limit of the 4U catalogue⁹. These three conditions all require that the cloud's temperature would have to exceed $\sim 10^8$ K. A cloud with this size, pressure and temperature could be confined by a mass $\geq 5 \times 10^6 M_\odot$. Future X-ray observations and a lower frequency VLB study could in principle increase this lower bound to a value such that pressure equilibrium would be unlikely.

Outer jet: comparing the VLB model with the 2.7 GHz map¹, we see that the angle of expansion of the jet has a constant value $\theta \sim 3^\circ$ as the radial distance, r , increases from ~ 1 pc to ~ 100 kpc. The minimum energy density meanwhile falls by a factor $\gtrsim 10^8$ from the nuclear value giving a minimum pressure $\sim 3 \times 10^{-12}$ dyn cm $^{-2}$. Unless NGC6251 is surrounded by a comparatively cool gaseous halo or lies in the cluster with which it is positionally identified^{1,10}, it is unlikely that the outer jet can be in pressure equilibrium with its surroundings.

Hot spots: the outer jet has not been detected beyond ~ 200 kpc but it points roughly towards the centroid of emission of the western component of the double. Two high brightness 'hot spots' are identifiable within this component, subtending an angle $\sim 40^\circ$ from the nucleus. It is tempting to identify these hot spots with either a vortex ring associated with an invisible extension of the outer jet or a precessional motion of the source of the jet, although the observed linearity of the jet renders this second explanation rather unlikely.

As with stronger sources¹¹, we can use the minimum pressure in the hot spots to put lower bounds on the momentum supply, P , and the velocity, V , of advance of the radio component through the surrounding medium¹²,

$$P \gtrsim 2 \times 10^{35} \text{ dyn}$$

$$V \gtrsim 10^8 \Omega_{\text{igm}}^{-1/2} \text{ cm s}^{-1}$$

measuring the surrounding density Ω_{igm} in units of the critical density $\sim 5 \times 10^{-30}$ g cm $^{-3}$.

Tail: there is also a low surface brightness 'tail', prominent at low frequencies, apparently in pressure equilibrium with its

surroundings. The intergalactic gas temperature T_{igm} then satisfies

$$T_{\text{igm}} \gtrsim 4 \times 10^7 \Omega_{\text{igm}}^{-1} \text{ K}$$

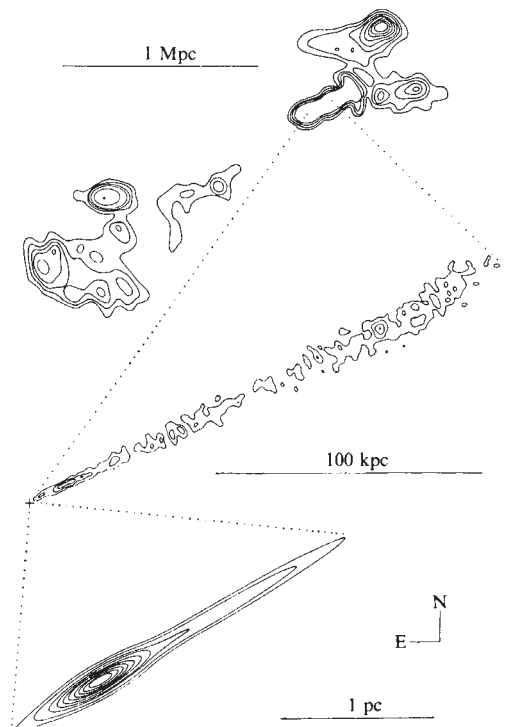
The total field and particle energy contained in the western component must exceed $\sim 10^{59}$ erg. If this is supplied at a constant rate over a time $0.5\tau/V$ (ref. 13), the power must satisfy

$$L \gtrsim 10^{43} \Omega_{\text{igm}}^{-1/2} \text{ erg s}^{-1}.$$

Discussion

There are now several known cases of sources in which the radio structure in the nucleus is aligned with larger scale structure¹⁴⁻¹⁷. In NGC6251, we see that the position angle is constant over a range $\gtrsim 5 \times 10^6$ in angular size (compared with $\gtrsim 3 \times 10^5$ in 3C111 (ref. 14)) and that the asymmetry of the outer jet is repeated in the nucleus. The observations reported here would seem to provide further support for theories of extragalactic radio sources in which mass, energy and momentum are supplied continuously to the radio components by a beam emanating from the nucleus of the associated galaxy¹⁸. Further insight into the nature of both compact and

Fig. 2 The radio source associated with NGC6251, showing the disposition of the nuclear and outer jets relative to the total source. The upper two contour maps are taken from Waggett *et al.*¹, the nuclear component has the contours of the adopted model described in the text.



extended radio sources may come from observing what variability, if any, is exhibited by the nucleus. For the remainder of this discussion we assume that all energy and momentum of the outer component have been supplied through both the nuclear and the outer jets.

Dynamics of the jet

A fairly strong case can be made for the flow velocity in the nuclear jet, v , being supersonic. If the flow were subsonic, then the pressure in the nuclear region $p \sim 4P/\pi d^2$ must exceed $\sim 3 \text{ dyn cm}^{-2}$, which by arguments similar to those used above is far too large to be confined.

A slightly weaker case can be made for the jet being free¹³ (not in pressure equilibrium with its surroundings) on the basis of the following arguments.

(1) The expansion angle θ seems to be constant over a range $\gtrsim 10^5$ in radius as might be expected in a free jet. If the jet were to remain in pressure equilibrium, a rather unlikely pressure variation, $p \propto r^{-1.7}$ would have to be maintained over this range.

(2) Comparing the equipartition field strengths in the nuclear and outer jets leads to a scaling $B \propto r^{-0.9}$, close to the $B \propto d^{-1} \propto r^{-1}$ variation expected for a free jet¹⁹. The magnetic field should be predominantly perpendicular to the jet leading to the prediction of a parallel electric vector (unaffected by Faraday rotation in the outer jet) in linear polarisation observations¹⁹. If the flow were isentropic, the particle pressure would fall more rapidly with r than the magnetic pressure. So, in order to prevent the plasma from becoming field-dominated and presumably unstable, dissipative processes like reconnection must continuously heat the particles, keeping their energy density comparable to that residing in the field. This is a justification of the equipartition hypothesis.

(3) Kelvin-Helmholtz instabilities, which may destroy a confined jet, should not be too important in a free jet^{20,21}.

(4) As discussed above, it may be difficult to supply the confining pressure for both the nuclear and the outer jet.

(5) The radio emission does not seem to come from the walls of the outer jet as might be expected if it were confined¹. Perpendicular polarisation would also be predicted if this were the case.

If the jet is free, then the Mach number $\mathcal{M} \gtrsim \theta^{-1} \sim 20$; otherwise the jet will seem to diverge more rapidly than is observed. Adopting the minimum pressure at $r \sim 100 \text{ kpc}$, this gives a momentum supply $P \gtrsim 3 \times 10^{35} \text{ dyn}$. If $\Omega_{\text{ig}} \lesssim 0.1$, then assuming $v \gtrsim 2V \gtrsim 10^9 \text{ cm s}^{-1}$ gives a minimum power $L \sim Pv/2 \gtrsim 10^{44} \text{ erg s}^{-1}$ and a mass loss $\dot{M} \sim 4(P/3 \times 10^{35} \text{ dyn})(v/10^9 \text{ cm s}^{-1})^{-1} M_{\odot} \text{ yr}^{-1}$, more than adequate to supply the minimum energy and mass requirements of the outer compo-

nent. If v (but not V) is mildly relativistic, then the only significant change in these estimates is that the surface brightness of a hypothetical eastern jet, similar in physical properties to the western jet, can be reduced to an unobservably low level, but only at the expense of increasing L to $\gtrsim 10^{45} \text{ erg s}^{-1}$.

In one possible mechanism for producing a collimated supersonic beam, a hot gas becomes trans-sonic by flowing through a de Laval nozzle^{19,22}. If this is the case in NGC6251, then we can estimate the stagnation pressure of the flow. Adopting the equipartition pressure in the nuclear jet, we estimate $\mathcal{M} \gtrsim 40$. For a specific heat ratio of magnetised plasma in the jet $\lesssim 1.5$, the stagnation pressure p_0 and the diameter of the nozzle d^* satisfy

$$p_0 \gtrsim 10^5 \text{ dyn cm}^{-2}, \\ d^* \sim 10^{15} (p_0/10^5 \text{ dyn cm}^{-2})^{-1/2} \text{ cm},$$

where $p_0 \sim 10^5 \text{ dyn cm}^{-2}$ requires isentropic flow for $r \lesssim 1.5 \text{ pc}$. It is difficult to confine and collimate the trans-sonic flow in a region of pressure $\sim p_0$ and size $\gtrsim 3d^*$, unless it is optically thick and radiation dominated—conditions that may exist around an accreting black hole of mass $\gtrsim 10^8 M_{\odot}$ (ref. 23).

We thank P. C. Waggett, P. J. Warner and J. E. Baldwin for use of their data before publication, and for allowing us to reproduce two of their maps in the composite picture of the source (Fig. 2). We are grateful to A. T. Moffet for help with the new 2.8 cm receiver at OVRO, and to G. A. Seielstad for assistance with the observations. We also thank W. L. W. Sargent and P. J. Young for useful discussions. This work is supported by NSF grants AST 77-00247, AST 76-01619 and AST 76-20375. The NRAO is operated by Associated Universities Inc. under NSF contract. Haystack Observatory is supported by NSF grant AST 14701.

Received 5 January; accepted 26 January 1978.

1. Waggett, P. C., Warner, P. J. & Baldwin, J. E. *Mon. Not. R. astr. Soc.* **181**, 465 (1977).
2. Clark, B. G. *Proc. I.E.E.E.* **61**, 1242 (1973).
3. Cohen, M. H. *et al. Nature* **368**, 405 (1977).
4. Baars, J. W. M. & Hartsuijker, A. P. *Astron. Astrophys.* **17**, 172 (1972).
5. Jennison, R. C. *Mon. Not. R. astr. Soc.* **118**, 276 (1958).
6. Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* (in the press).
7. Burbidge, G. R., Jones, T. W. & O'Dell *Astrophys. J.* **193**, 43 (1974).
8. Scott, M. A. & Readhead, A. C. S. *Mon. Not. R. astr. Soc.* **180**, 539 (1977).
9. Foreman, W. *et al. Astrophys. J. Suppl.* (in the press).
10. Zwicky, F. *Catalogue of Galaxies and of Clusters of Galaxies*, (California Institute of Technology, 1961).
11. Hargrave, P. J. & Ryle, M. *Mon. Not. R. astr. Soc.* **166**, 305 (1976).
12. Turland, B. D. *Mon. Not. R. astr. Soc.* **172**, 181 (1975).
13. Scheuer, P. A. G. *Mon. Not. R. astr. Soc.* **166**, 513 (1974).
14. Kellerman, K. I., Clark, B. G., Niell, A. E. & Shaffer, D. B. *Astrophys. J. Lett.* **197**, L113 (1975).
15. Fomalont, E. B. & Miley, G. K. *Nature* **257**, 99 (1975).
16. Pauliny-Toth, I. I. K. *et al. Nature* **259**, 17 (1976).
17. Wilkinson, P. N., Readhead, A. C. S., Purcell, G. H. & Anderson, B. *Nature* **269**, 764 (1977).
18. De Young, D. S. *A. Rev. Astron. Astrophys.* **14**, 447 (1976).
19. Blandford, R. D. & Rees, M. H. *Mon. Not. R. astr. Soc.* **169**, 395 (1974).
20. Turland, B. D. & Scheuer, P. A. G. *Mon. Not. R. astr. Soc.* **176**, 411 (1976).
21. Blandford, R. D. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **169**, 443 (1976).
22. Wiita, P. J. *Astrophys. J.* (in the press).
23. Rees, M. J. *Proc. 8th Texas Symp. on Relativistic Astrophysics* (1977).

³¹P magnetic resonance of DNA in nucleosome core particles of chromatin

N. R. Kallenbach & D. W. Appleby

Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

C. H. Bradley

Bruker Instruments Inc., Manning Park, Billerica, Massachusetts 01821

The ³¹P magnetic resonance spectrum of nucleosome cores from calf thymus nuclei has been measured. The spectrum consists of a single symmetric line, with a maximum second component of 5%. It is argued that this reflects a single average state of the DNA backbone, and does not support models that invoke backbone kinking at a frequency of 1 in 10 base pairs.

THE existence of a repeating substructure in chromatin is by now well documented¹. There are two components to the overall repeating unit: the nucleosome core particle, or nu body, consisting of two molecules each of four histones, H2A, H2B, H3 and H4, complexed with a length of DNA of 140 base pairs²⁻⁴, and an additional variable length of DNA 20–60 base pairs long⁵ probably associated with the histone H1 (ref. 6). Structural studies on the nucleosome core particles

based on electron microscopy^{7,8} and X-ray⁹ and neutron diffraction^{10,11} define a deformed cylindrical or spherical unit. In crystals the particle appears wedge shaped with dimensions $57 \times 110 \times 110 \text{ \AA}$ (ref. 12). From neutron-diffraction data, the DNA molecule can be assigned a radius of gyration of $50 \text{ \AA}$, whereas the radius of gyration of the proteins is $30 \text{ \AA}$ (ref. 10). Therefore, the DNA lies towards the outside of the particle, not tightly wound at the centre¹⁰. This is consistent with the striking regularity in size distribution of the DNA obtained from nucleosomes following nuclease digestion. One site per 10 base pairs seems to be highly susceptible to attack¹³⁻¹⁶, and presumably these sites lie at or very near the surface of the particle.

There is little solid information concerning the precise folding of the DNA within nucleosome cores. It is clear that some folding of DNA is necessary to accommodate 140 base pairs of DNA B helix (nearly $500 \text{ \AA}$) into a core particle of the dimensions indicated. Formation of a core introduces approximately one left-handed superhelical twist into covalently closed circular DNA^{17,18}. Two extreme models for the packing of DNA in cores have been proposed. In one, the double helix is kinked at regular intervals so as to permit sharp bending of the helix axis¹⁹⁻²¹. In the other, the helical axis is altered gradually and uniformly so as to produce a smoothly turning helical axis or DNA backbone, the strain being distributed over all base pairs in the molecule (ref. 12 and J. L. Sussman & E. N. Trifanov, personal communication). Between these extremes, there can exist numerous intermediate models, so that structural studies at much higher resolution than now available are needed to establish the state of the DNA backbone in nucleosome cores.

To discriminate between the two extreme models we investigated directly the state of the DNA backbone in nucleosome cores by measuring the ^{31}P NMR spectrum of cores prepared by micrococcal nuclease digestion of nuclei (here isolated from calf thymus glands). There are two strong arguments for ^{31}P NMR spectroscopy of nucleosomes: first, there is only one phosphate per DNA nucleotide in contrast to the multiplicity of protons, for example, in both the DNA and histones. Second, there is now considerable evidence derived from ^{31}P NMR data on mononucleotides^{22,27}, dinucleotides^{23,24}, oligo- and polynucleotides²⁵⁻²⁷ and intercalation complexes of these^{25,28} that the ^{31}P chemical shifts in phosphate esters reflect the local conformation of the phosphate group²⁹. The P-O5' and O3'-P torsional angles constitute the source of most conformational freedom in a nucleotide unit³⁰, and any substantive kinking involving alteration of these angles should be accompanied by correspondingly distinctive ^{31}P chemical shifts, longitudinal relaxation times or both. The proton decoupled ^{31}P NMR spectrum of fractionated monosomes consists of a single symmetric peak at the highest resolution we have obtained. This suggests that the phosphodiester backbones of DNA within nucleosome cores correspond to a single conformational state, not two as would be required for kinking with a frequency of 1 nucleotide per 10 to 20 nucleotides. We find some indication that the ^{31}P chemical shift of nucleosome cores is shifted very slightly upfield from that of the DNA extracted from the particles. This is difficult to interpret, but is not inconsistent with a picture in which each base pair is strained slightly in packing the DNA within a nucleosome.

Preparation and characterisation of nucleosome cores

Nuclei were prepared from fresh or frozen calf thymus glands according to a modification of the method of Marshall and Burgoyne³¹ which minimises endogenous nuclease activity. Briefly, 100 g of gland tissue was minced and homogenised using a Dounce homogeniser in buffer HB (340 mM sucrose, 60 mM KCl, 15 mM NaCl, 0.5 mM spermidine, 0.15 mM spermine (Calbiochem), 15 mM Tris, pH 7.4, 2 mM EDTA, 0.5 mM EGTA, 15 mM mercaptoethanol and 1 mM phenyl methyl sulphonyl fluoride (PMSF, Calbiochem)). The homogenate was filtered through eight layers of cheesecloth and

rehomogenised. Nuclei were pelleted by centrifugation at 200g for 10 min, washed and recentrifuged in a second buffer, RB, identical to HB except that EDTA and EGTA concentrations are 0.1 mM. The final white pellet was resuspended in 200 ml of this buffer, 400 ml glycerol was added and it was stored frozen until required.

Nuclei were digested with micrococcal nuclease (Worthington) and lysed in hypotonic buffer containing deoxycholate and Triton X-100. This process is summarised in Table 1. The final supernatant, S IV was layered on a 5–18% nonlinear sucrose gradient (made as described by Noll³² and centrifuged as described in Fig. 1). Fractions from these tubes corresponding to nucleosome monomers were pooled, dialysed exhaustively against 0.1 mM EDTA, 0.1 mM EGTA, 1 mM Tris pH 7.4 and concentrated to 1.88 ml for NMR spectroscopy.

Proteins from this material were analysed on sodium dodecyl sulphate (SDS)–polyacrylamide gels, according to the method of Laemmli³³. The DNA from the nucleosomes was analysed on mixed agarose–polyacrylamide gels³⁴ and stained with $1 \mu\text{g ml}^{-1}$ ethidium bromide. Such a gel is shown in Fig. 2. Molecular weight markers used were a *HincII* digest of ΦX174 RF DNA obtained from New England Biolabs.

^{31}P NMR of nucleosome cores and DNA extracted from cores

The (145.7 MHz) ^{31}P NMR spectrum of the fractionated monomer preparation is shown in Fig. 3, measured using Fourier transform spectroscopy on a Bruker WH360 spectro-

Table 1 Preparation of nucleosome suspensions

Sample	% Acid soluble	Total acid precipitable (A_{260} units)
t_0	0	3,480
t_8	29	2,490
SI	29	140
SIH	0	140
SIH	0	140
SIV	0	1,440
PEL	0	630

Two 40 ml samples of nuclear suspension prepared as described in the text were diluted with 160 ml each of RB. The nuclei were centrifuged at 2,000 r.p.m. for 5 min in the GSA rotor (Sorvall). Each nuclear pellet was resuspended in 40 ml of RB and 0.2 ml of 1M CaCl_2 were added with shaking at 37°C (New Brunswick rotary shaker). After 5 min 40 μl were dissolved in 1.0 ml of 1% SDS, 5M urea and 40 μl were precipitated in 1M NaCl, 1M perchloric acid (PCA). These samples represent t_0 . After a minimum of 15 min on ice the PCA sample was centrifuged (10,000g, 2 min) and A_{265} for the supernatant was read on a Cary 16 spectrophotometer. (This value was between 0 and 0.010). On taking t_0 , 1,500 units of micrococcal nuclease (Worthington) were added. Time points of 40 μl were taken into 1.0 ml of 1M NaCl, 1M PCA and treated as for t_0 . The A_{265} values were divided by 1.625 to correct for hyperchromism. In this way a plateau value of $29 \pm 2\%$ acid soluble A_{265} units was obtained. t_8 was 1 h and 2 h respectively for the two preparations. Plateau was reached in approximately 45 min in both preparations. After taking t_8 each reaction was stopped by adding 2.5 ml of 0.1M EGTA and the flask was cooled on ice. The nuclei were pelleted at 5,000 r.p.m. for 5 min (HB-4 rotor, Sorvall). The supernatant was retained as SI. The pellet was then washed and lysed as follows. The pellet was resuspended in 20 ml of 0.1mM EDTA, 0.1mM EGTA, 0.1% NP-40 and left on ice for 15 min. The nuclei were pelleted as above to produce supernatant SIH. The pellet was resuspended in 20 ml of 0.1 mM EDTA, 0.1mM EGTA, 0.1% Triton X-100 (TX-100). The nuclei were left on ice for 15 min. Again they were pelleted as above to produce supernatant SIIH. The pellet was lysed by resuspension in 10 ml 0.1mM EDTA, 0.1mM EGTA, 0.1% TX-100. Both preparations were combined. Solid deoxycholate was added to 0.7% for 5 min on ice and the nuclei were immediately centrifuged at 10,000 r.p.m. for 5 min to produce supernatant SIV and PEL. Each of these supernatants and the pellet was analysed for the amount of acid soluble and total A_{260} units it contained. The results are shown above as combined values for both preparations. The yield of A_{260} units in SIV was 58% of the total acid precipitable units present at t_8 (41% of the total present at t_0).

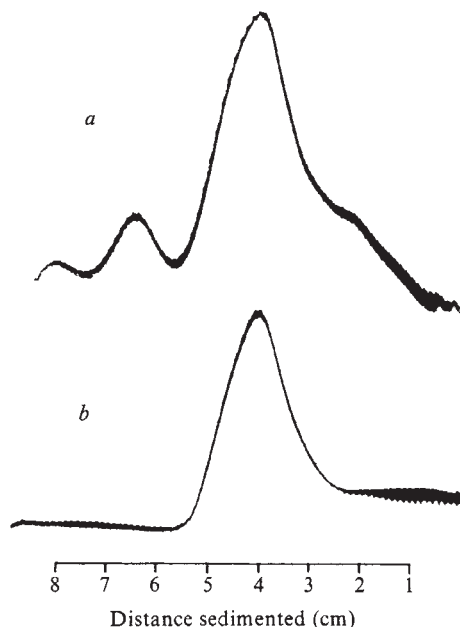


Fig. 1 *a*, SIV 7.5 ml (72 A_{260} units per ml) was centrifuged in 0.35 ml aliquotes on exponential sucrose gradients prepared as described by Noll³² ($C_r = 27\%$ w/v, $V_m = 9.428$ ml, $\rho = 1.51$) for the Beckman SW41 rotor. Centrifugation was carried out at 40,000 r.p.m. for 15–18 h at 4 °C. The tubes were fractionated from the bottoms and A_{260} was continuously monitored with a Gilford flow cell and a Gilford spectrophotometer. The absorbance output from a representative gradient is shown. The major peak was taken (11–12S) from each gradient, combined, dialysed against 0.1 mM EDTA, 0.1 mM EGTA, 1 mM Tris pH 7.4 and concentrated to 1.88 ml. The A_{260} of the resulting monosome preparation was 163. *b*, 21 μ l of this solution was analysed on another gradient in the same conditions as described above to determine the degree of disome contamination. It may be seen that the monosome preparation is free of higher order oligomers.

meter. Since there is no obvious indication of two components in the spectrum, we carried out a digital subtraction of the spectrum reversed right to left from this spectrum about the maximum of the peak. For a symmetrical resonance $f(\nu)$ symmetrical about $\nu = \nu_0$, this process constructs values $f(\nu_0 + \nu') - f(\nu_0 - \nu') = 0$, whereas any contribution at a frequency different from ν_0 should give rise to nonzero values of this difference, positive and negative lobes of the form seen in Fig. 3*b*. In carrying out this procedure the phasing of the peak is critical, and some part of the difference spectrum in Fig. 3*b* may be ascribed to imperfect phasing of the broad line. The observed difference, 5.5% of the amplitude of the peak maximum, thus represents an upper limit and the actual case may be considerably smaller. In other spectra and samples we have seen an asymmetry below 3%.

A second possibility is that despite the lack of a distinct second component due to phosphodiester linkages in a particular state or environment, the spin lattice relaxation times (T_1) of the bulk and minor component(s) might differ. The T_1 relaxation time of the nucleosome core preparation was determined by the inversion recovery method using 180° – t – 90° pulse sequences on the Bruker WH360 spectrometer at 145.7 MHz. The major component in the ^{31}P spectrum exhibits a single relaxation time $T_1 = 2.4$ s, with no apparent second component with any distinct T_1 values indicated, as for example can be detected in the case of yeast tRNA^{Phe} (ref. 27). We also remeasured the spectrum in conditions in which the major resonance of value $T_1 = 2.4$ s is suppressed (see ref. 27); again no shift or clear asymmetry of the resulting spectrum could be seen.

Comparison of the ^{31}P NMR spectrum of the DNA in cores after proteolytic digestion of the histones is shown in Fig. 4. Following the digestion and restoring the same pH, there is

observed a slight difference in chemical shift (6 Hz at 24.28 MHz), and a pronounced sharpening of the resonance (see ref. 38). It is worth noting that the line width increases by a factor of about 11 with a frequency increase by a factor of 6, similar in ratio to that observed for ^{31}P resonances in tRNA²⁶. Comparison of two spectra at 109 MHz (Bruker HX270 spectrometer) and 145.7 MHz show a ratio of line widths of 1.37 compared to a ratio of frequencies of 1.33, suggesting that the relaxation mechanism in nucleosome cores is not exclusively dominated by the anisotropic chemical shift²⁶.

The DNA in nucleosome cores

In order to interpret the finding that the maximum asymmetry in the ^{31}P NMR spectrum from calf thymus nucleosome cores is $\leq 6\%$, it is necessary to establish a basis for asserting that the presence of 1 kink in 10 base pairs would give rise to a distinct ^{31}P chemical shift from the remaining unknicked phosphodiester bonds. The argument hinges on two premises: (1) in mono- and oligonucleotide model systems, ^{31}P chemical shifts provide a sensitive measure of conformation about the O–P–O bond, and (2) it is unlikely that the kinked residues rapidly interconvert with unknicked ones, which could lead to coalescence of the distinct resonances from each state.

Surveys of the chemical shifts in mononucleotides, dinucleotides and oligonucleotides indicate the following progression

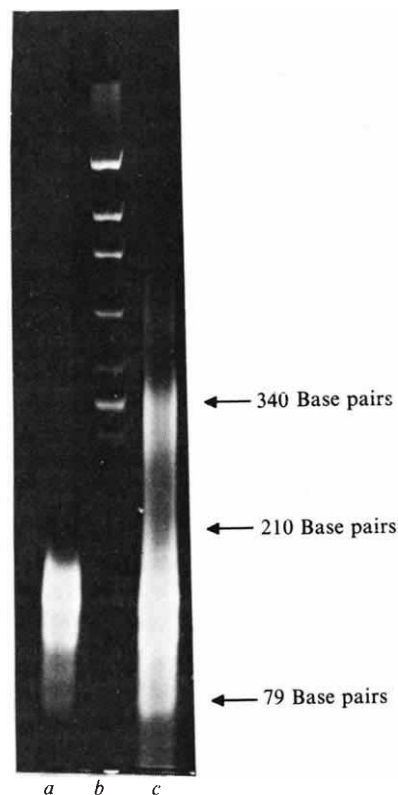


Fig. 2 Agarose-polyacrylamide gel electrophoresis of DNA from monosome preparation. 1 μ l of the proteinase K-digested nucleosomes described in the legend to Fig. 4*c* was added to 10 μ l of Tris-borate running buffer³⁴. Then 1 μ l glycerol and 1 μ l of 2M KCl were added. After 15 min on ice the potassium precipitate of decyl sulphate was pelleted at 10,000*g* for 2 min and the supernatant was loaded in slot *a* of the gel made as described in ref. 34. DNA (5 μ g) obtained by lysing and phenol extracting the pellet—see Table 1—was loaded in slot *c*. *Hinc*II restriction fragments of Φ X174 RF DNA (New England BioLabs) (4 μ g) were loaded in slot *b*. Electrophoresis was carried out at 100 V, 40 mA for a time which placed the bromophenol blue dye (loaded in an adjacent slot) 1 cm from the base of the gel. The gel was removed and stained for 30 min in 1 μ g ml⁻¹ ethidium bromide in 10 mM Tris pH 7.4, 10 mM NaCl. It was photographed on Polaroid type 55, 4 $\times$ 5 direct positive film using ultraviolet illumination and a red filter.

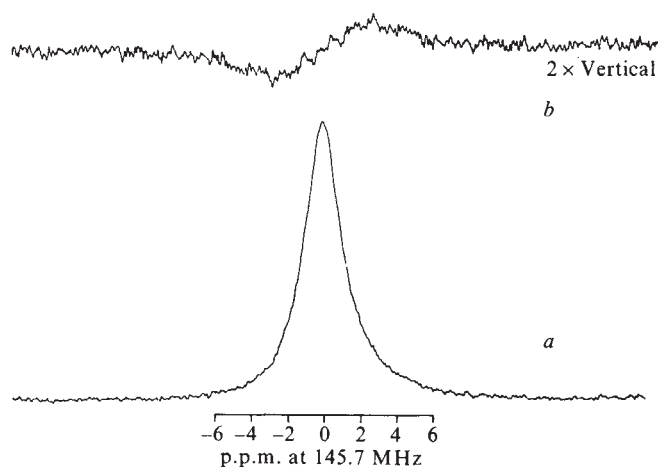


Fig. 3 *a*, ^{31}P NMR spectrum of nucleosome cores. A sample of 1.88 ml (163 A_{260} per ml) in 0.1 mM EDTA, 0.1 mM EGTA, 1 mM Tris pH 7.4 was diluted with 0.2 volumes D_2O in a 10 mm tube. The Fourier transform spectrum at 145.7 MHz was measured on a Bruker WH360 spectrometer; 6,500 transients, with an acquisition time of 0.4 s and a delay of 9 s. Temperature was controlled at $+5^\circ\text{C}$. *b*, The result of subtracting the right to left reversed spectrum from spectrum *a*, recorded at twice the vertical scale.

in ^{31}P chemical shifts at 17°C and neutral pH: 3', 5' cyclic nucleotides, $+1.5$ to 2.1 p.p.m. (from 85% phosphoric acid as a reference); dinucleotides and polynucleotides, $+0.5$ to 1.1 p.p.m.; 3'-mononucleotides, -0.4 to 0.2 p.p.m.; and 2', 3' cyclic nucleotides, below -20 p.p.m. (ref. 23). The highly strained 2', 3' cyclic nucleotide structure actually exhibits a

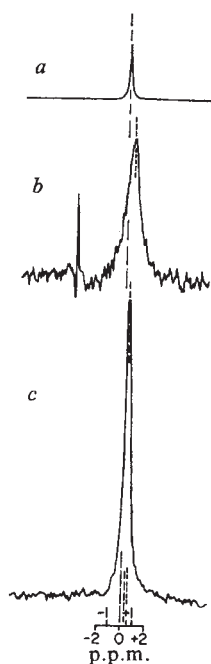


Fig. 4 ^{31}P NMR spectrum of cores and DNA from cores. The same sample as in Fig. 3 was measured, at 15°C , on a Varian NV14 spectrometer at 24.28 MHz (*b*). Conditions were: 0.8 s acquisition time, 5 s delay, 9,269 transients. Panel *a* shows a reference phosphoric acid solution at $+1.6$ p.p.m. from standard 85% phosphoric acid in D_2O . (The peak at the left in panel *b* is due to the transmitter, and is not from sample.) Panel *c* is the spectrum of a proteinase K digest of 1.3 ml of the sample in *b*. To this volume was added 0.13 ml 10% SDS, 0.13 ml 5M NaCl 0.13 ml 5 mg ml^{-1} proteinase K (EM), 1 ml H_2O and the solution was incubated for 1 h at 37°C . A second volume of 0.013 ml proteinase K was added, and incubated for a further 1 h. The mixture was concentrated to 1 ml, the pH adjusted to 7.4 with 10 μl of 1M Tris, and the spectrum measured as before (10,354 transients).

deformation in the O-P-O bond angle³⁵. Dinucleoside monophosphates exhibit shifts in ^{31}P resonance from 0.2 to 0.6 p.p.m. downfield as temperature increases and unstacking occurs²⁴.

Intercalation of actinomycin D into minihelices of dpGpC leads to a ^{31}P chemical shift of the bridge phosphate of -1.7 to -2.4 p.p.m. (ref. 28). A similar shift is observed in the 1:2 complex between actinomycin D and dApTpGpCpApT (ref. 36). Other molecules and dinucleotides exhibit differential shifts between -0.1 p.p.m. and $+0.2$ p.p.m. depending on temperature and stoichiometry²⁸. Thus the ^{31}P chemical shift is sensitive to the process of intercalation, and generally to the conformation of the O-P-O backbone. Gorenstein and his coworkers have argued that gauche phosphate ester conformations are generally associated with higher field shifts than trans or other nongauche states^{22,24,29}. Provided that the postulated kink is associated with an alteration of the backbone conformation, as in the case of the kind of model proposed by Sobell *et al.*^{20,21} for example, there should be an equilibrium population of 10% distinct ^{31}P chemical shifts.

This population need not correspond exactly to the intensity of the ^{31}P resonance measured in conditions of full proton decoupling, however. If the two states interact differently with protons, there could arise a relative distortion of the peak intensities due to nuclear Overhauser effects²⁷. We have repeated our measurements in the absence of proton decoupling, and find essentially the same profile. Hence there is no basis to argue that there is a selective loss in the resonance from a particular class within the population of phosphodiester bonds.

The second premise concerns the possibility that two distinct resonances are present but their rapid interconversion leads to a single coalesced resonance. For this kind of exchange effect to occur, if we suppose the difference in frequency between the two peaks to correspond to 0.5 p.p.m. at 109 MHz, the exchange rate would have to be roughly $\pi/\sqrt{2} (\Delta\nu)$ or about 10 s^{-1} (see ref. 37, for example). Recent experiments in which the DNA of the nucleosome core was end-labelled with ^{32}P and the absolute size of the DNA resulting from nuclease treatment was measured imply a precise and stable selection of sites for attack along the DNA¹⁵. This mitigates against any rapid shuttling or interchange between kinked and unkinked residues; it implies a static population of vulnerable sites for nuclease attack determined by the overall structure of the core particle. Similarly the non-uniformity of rates of nuclease attack at different positions of the DNA suggests that no rapid exchange process between kinked sites with periodicity of 10 base pairs on DNA is likely to occur.

There are then two conclusions from this study. First, unless kinking occurs by a mechanism that leaves the phosphodiester bonds unperturbed, a maximum non-uniformity of 1 bond in 20 occurs, so that kinks as frequent as 1 in 10 of the kind postulated by Sobell *et al.*^{20,21} can be excluded. The model proposed by Crick and Klug¹⁹, however, works by a large alteration in the C'₄-C'₅ torsion angle, with little change in the P-O₅' or O₃'-P angles. It conceivably thus might not lead to an alteration of the ^{31}P chemical shift of the adjacent phosphodiester bonds. This type of structure cannot be excluded, since there are no ^{31}P NMR data giving chemical shifts in compounds differing exclusively in the C'₄-C'₅ torsional angle. Second, several models for kinks become allowed if their frequency is 1 per 20 base pairs or less; this still permits considerable compacting of the DNA in nucleosome cores. We are attempting to establish this limit more precisely, and to determine the status of phosphodiester bonds in the spacer DNA between cores, in the presence or absence of histone H1.

This work was supported by American Cancer Society grant NP-187A. We thank Rong Ine Ma for technical assistance, and Dr David Montgomerie for help in our centrifugation work.

Received 31 October 1977; accepted 9 January 1978.

- Weintraub, H., Worcel, A. & Alberts, B. *Cell* **9**, 409-417 (1976).
- Sollner-Webb, B. & Felsenfeld, G. *Biochemistry* **14**, 2915-2920 (1975).
- Axel, R. *Biochemistry* **14**, 2921-2925 (1975).

4. Shaw, B. R., Herman, T. M., Kovacic, R. T., Beaudreau, G. S. & Van Holde, K. E. *Proc. natn. Acad. Sci. U.S.A.* **73**, 505–509 (1976).
5. Lohr, D., Corden, J., Tatchell, K., Kovacic, R. T. & Van Holde, K. E., *Proc. natn. Acad. Sci. U.S.A.* **74**, 79–83 (1977).
6. Noll, M. & Kornberg, R. D. *J. molec. Biol.* **109**, 393–404 (1977).
7. Woodcock, C. L. F. *J. Cell Biol.* **59**, 368a (1973).
8. Olins, A. L. & Olins, O. E. *Science* **183**, 330–332 (1974).
9. Pardon, J. F. & Wilkins, M. H. F. *J. molec. Biol.* **68**, 115–124 (1972).
10. Pardon, J. F. *et al. Nucleic Acids Res.* **2**, 2163–2176 (1975).
11. Baldwin, J. P., Boseley, P. G., Bradbury, M. & Ibel, K. *Nature* **253**, 245–249 (1975).
12. Finch, J. T. *et al. Nature* **269**, 29–36 (1977).
13. Noll, M. *Nucleic Acids Res.* **1**, 1573–1578 (1974).
14. Simpson, R. T. & Whitlock, J. P., Jr *Nucleic Acids Res.* **3**, 117–127 (1976).
15. Whitlock, J. P., Rushizky, G. W. & Simpson, R. T. *J. biol. Chem.* **252**, 3003–3006 (1977).
16. Sollner-Webb, B. & Felsenfeld, G. *Cell* **10**, 537–547 (1977).
17. Griffith, J. D. *Science* **187**, 1202–1203 (1975).
18. Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843–1847 (1975).
19. Crick, F. H. C. & Klug, A. *Nature* **255**, 530–533 (1975).
20. Sobell, H. M., Tsai, C.-C., Gilbert, S. G., Jain, S. C. & Sakore, T. D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3068–3072 (1976).
21. Sobell, H. M., Tsai, C.-C., Jain, S. C. & Gilbert, S. G. *J. molec. Biol.* **114**, 333–365 (1977).
22. Gorenstein, D. G. *J. Am. chem. Soc.* **97**, 898–900 (1975).
23. Cozzzone, P. & Jardetzky, O. *Biochemistry* **15**, 4860–4865 (1976).
24. Gorenstein, D. G., Findlay, J. B., Momii, R. K., Luxon, B. A. & Kar, D. *Biochemistry* **15**, 3796–3803 (1976).
25. Patel, D. *Biopolymers* **15**, 533–558 (1976).
26. Gueron, M. & Shulman, R. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3482–3485 (1975).
27. Hayashi, F., Akasaka, K. & Hatano, H. *Biopolymers* **16**, 655–677 (1977).
28. Reinhardt, C. G. & Krugh, T. R. *Biochemistry* **16**, 2890–2895 (1977).
29. Gorenstein, D. G. & Kar, D. *Biochem. biophys. Res. Comm.* **65**, 1073–1080 (1975).
30. Kallenbach, N. R. & Berman, H. M. *Q. Rev. Biophys.* **10**, 137–236 (1977).
31. Marshall, A. J. & Burgoyne, L. A., *Nucleic Acid Res.* **3**, 1101–1110 (1976).
32. Noll, H. *Nature* **215**, 360–363 (1967).
33. Laemmli, U. K. *Nature* **222**, 680–685 (1970).
34. Peacock, A. C. & Dingman, C. W. *Biochemistry* **7**, 668–674 (1968).
35. Coulter, C. L. & Greaves, M. L. *Science* **169**, 1097–1098 (1970).
36. Patel, D. J. *Biochemistry* **13**, 2396–2402 (1974).
37. Carrington, A. & McLachlan, A. D. *Introduction to Magnetic Resonance* (Harper and Row, London, 1967).
38. Hanlon, S., Glonek, T. & Chan, A. *Biochemistry* **15**, 3869–3875 (1976).

Neurophysiological basis of directional hearing in amphibia

Alan Pettigrew, Shin-Ho Chung & Michael Anson

National Institute for Medical Research, London NW7, UK

Discrete regions of neural tissue along the anterior-posterior axis of the torus semicircularis act as resonators, each one broadly tuned to a different range of frequencies, and each one responding optimally to the incident sound from a different direction. In this way, a map of auditory space is represented in the midbrain.

Of the many auditory phenomena, directional hearing presents a special problem for neurobiologists. The analogous problem in the other senses, such as touch and sight, is less puzzling, since the stimulus excites a discrete subset of receptors, all of which project in a topographical order on to the central nervous system. With hearing, however, a sound excites both ears, and the only possible cues for directional hearing, as shown by psychophysical studies on humans^{1–3}, are very small differences in the time of arrival and intensity of sound at the two ears. For a small animal, such as the frog, the binaural difference in intensity becomes negligible, since the wavelengths of its audible frequencies^{4,5} are long and the head is only a small obstacle for producing an acoustic shadow, especially if the sound comes from a distant source. Thus, in determining directionality, the animal has to rely on a maximum of 40 μ s difference in the time of arrival of the sound at the two ears; yet, it can normally locate the source of the sound it hears with great accuracy⁶. How these cues are translated into the perception of directional hearing is unknown, but almost all the proposed theories^{7–11} relegate the entire task of analysis to the nervous system.

In this and the following article, we briefly report the results of investigations aimed at elucidating this question. We show that the direction of the incident sound is accurately represented in the auditory centre of the midbrain. This mapping of the auditory space is achieved by arranging the incoming fibres according to their frequency sensitivity, so that each midbrain region along the anterior-posterior axis acts like a broadly tuned resonator. Each resonator, in turn, is optimally tuned to the incident sound originating from a specific part of auditory space. These findings prompted us to enquire into the physical basis of directional hearing. Using the technique of laser speckle interferometry, we demonstrate¹² that the two eardrums are elastically coupled, and their motion is modulated by the resonance characteristics of the mouth cavity. Moreover, we show that the movement of the eardrums forms a physical basis for directional hearing. In addition, we provide all the

numerical constants concerning the mechanical properties of the amphibian middle ear.

Neural responses in the midbrain

The midbrain of *Rana temporaria* (or sometimes, *Rana esculenta*) was exposed, and the potentials evoked by a brief

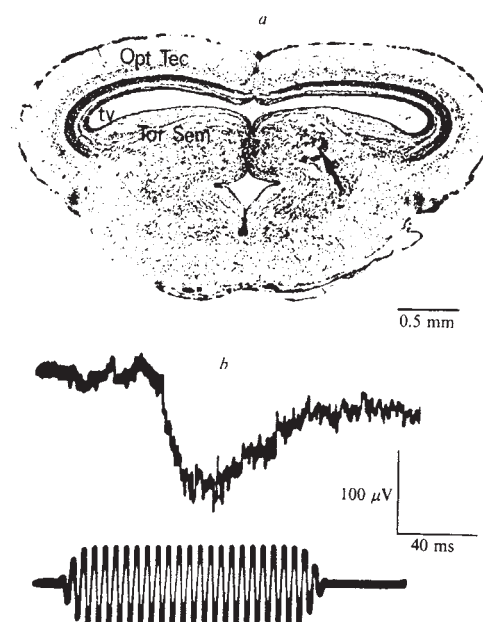


Fig. 1 Evoked auditory responses. *a*, After each experiment, the recording site in the torus semicircularis was identified by making an electrolytic lesion with the recording electrode (arrow). The section, cut in transverse plane, is stained with cresyl fast violet. Abbreviations: Opt Tec, optic tectum; Tor Sem, torus semicircularis; t. v., tectal ventricle. *b*, A brief burst of pure tone elicited a pronounced negative wave in the torus semicircularis (upper trace). The impedance of the recording electrode, when measured in normal saline at 10 kHz, was about 100 k Ω ; negative downward. The stimulus shown in the bottom trace is a pure tone of 200 Hz and was presented at a pressure level of 0.5 Pa. In all experiments, the sound pressure was continuously monitored with a pressure transducer (AKG CK2) suspended just above the ear, and the amplified output was expressed as root-mean-square.

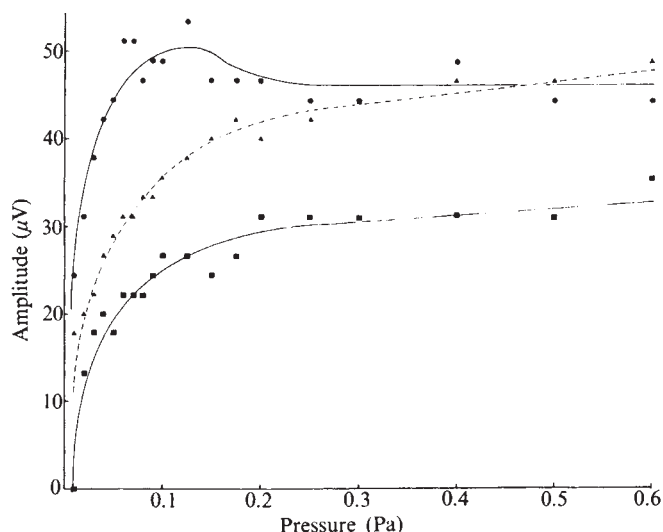


Fig. 2 Strength-response curves of the evoked potentials. As the sound pressure was increased, the amplitude of the response first increased rapidly to reach its maximum and remained stable thereafter. Note that the shape of the strength-response curves differs for different frequencies. For low frequencies (~ 200 Hz), not shown here, the amplitude of responses increases with pressure up to 0.5 Pa. Each point in this and subsequent figures represents an average of 8 responses. ●, 1,250 Hz; ▲, 1,500 Hz; ■, 2,000 Hz.

burst of pure tone were monitored in the torus semicircularis, which lies immediately below the tectum. When the pressure of the sound was 0.1 Pa (1 Pa ≈ 94 dB, where 0 dB = 2×10^{-5} N m $^{-2}$), the peak amplitude of these potentials was of the order of 50 μ V. To enhance the signal-to-noise ratio, we recorded the potentials differentially, using a matched pair of Dowben-Rose electrodes¹³. The junction potential between the tip of each electrode and the brain tissue was shunted with a 10 M Ω resistor connected between the gate and source of the input transistor. The amplified signals were then fed through a low-pass filter with a cut-off frequency of 5 kHz and the evoked responses, when needed, were averaged with a signal averager (Biomac 1000). The shape of the evoked potentials so recorded is slightly distorted by the impedance characteristics of the electrode, but this distortion does not affect our analysis significantly. All the figures illustrated in this and the following article were obtained in anechoic conditions.

After a brief auditory stimulus, a pronounced negative wave appears in the deep layer of the tectum. As the recording electrode penetrates the torus semicircularis, the amplitude of the potentials increases, reaching a maximum at about 100 μ m below the roof of the torus semicircularis (Fig. 1). We note here that the auditory nerves first terminate in hindbrain nuclei, from which the torus semicircularis receives its projection. These monophasic potentials, the peak of which occurs at 40 msec to 60 msec after the onset of the stimulus, represent the summed extracellular current generated by the activation of the presynaptic and postsynaptic neurones¹⁴. Evidence for this is based on detailed electrophysiological and pharmacological tests, which will be described elsewhere.

At any given electrode position along the anterior-posterior axis of the midbrain, the amplitude of the potential recorded is dependent on both the driving force and its frequency (Fig. 2). With sounds of small amplitude, the response of the auditory centre is essentially curvilinear. As the intensity is increased and the auditory mechanism is forced to vibrate with larger amplitude, the neural responses reach a plateau, implying that some portion of the mechanism reaches a constraining limit. Since we show¹² that the middle ear faithfully obeys Hooke's law, this distortion is brought about by the action of the inner ear or neural elements or both.

Spatial analyses of the evoked potentials indicate that the auditory centre in the midbrain is organised tonotopically. Fibres responding to sounds of high frequency project onto the rostral region of the torus semicircularis, whereas those responding to sounds of low frequency project to its caudal region. While keeping the sound pressure constant and below the level of neural saturation, we have measured the amplitude of responses for each of 35 different sound frequencies sequentially along the anterior-posterior axis of the midbrain. In this way, we have obtained a curve representing the amplitude of responses as a function of frequency for each midbrain position. As shown in Fig. 3, the upper limit of frequency response in the rostral midbrain is in the order of 3,500 Hz, whereas that of the caudal midbrain is approximately 1,300 Hz. Intermediate positions in the midbrain respond to intermediate tones. Although the analysis of discharge rates in single midbrain neurones may lead to the same conclusion, this will require a large sample of cells at each midbrain position and statistical analysis of data obtained from different animals. We believe that the pattern of the auditory projection is revealed more clearly and easily when population responses are studied¹⁵.

As in other sensory systems, the projection of the auditory inputs to the midbrain is predominantly crossed, as shown by experiments involving unilateral section of the auditory nerve. If the measurements described above are repeated after the

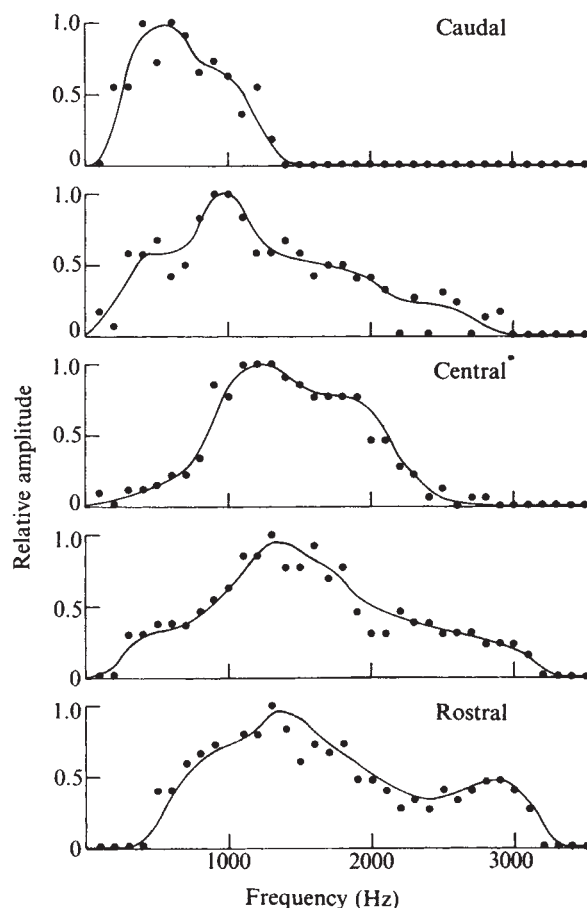


Fig. 3 Tonotopic organisation. With the recording electrode first placed in the caudal region of the torus semicircularis, the evoked responses at each of 35 different frequencies were averaged over 8 presentations of the stimulus. The relative amplitude of these responses are plotted against frequency (top). This process was then repeated after placing the recording electrode further rostrally in steps of 200 μ m. The sound pressure was kept constant throughout at 0.02 Pa. Note that the peak sensitivity shifts progressively towards high frequencies as the recording electrode is moved rostrally. The rostral end of the midbrain responds to high frequencies as well as to intermediate frequencies (bottom).

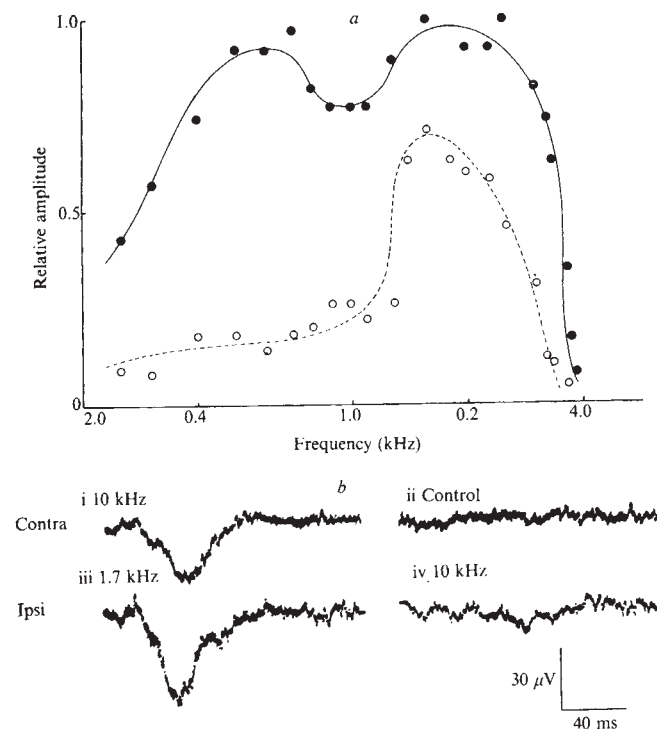


Fig. 4 Pattern of auditory projection. *a*, The amplitude of responses in the rostral area of the torus semicircularis is plotted against frequency of stimulus (log scale) when the contralateral auditory nerve was intact (●) or cut (○). There is an ipsilateral projection from each ear to the rostral region of the midbrain, and this projection conveys only high frequency responses. *b*, Evoked responses in the torus semicircularis were obtained by vibrating the eardrum with a piezoelectric crystal. The sinusoidal field applied across the crystal had the same form as that shown for the sound stimulus in Fig. 1*b*. Traces represent the averaged response to 8 presentations of a brief stimulus. Neural responses could be recorded when the contralateral eardrum was moved at 10 kHz (i). Control trace (ii) was obtained while the crystal was vibrating but not actually displacing the eardrum. Responses could also be recorded when the ipsilateral eardrum was moved at frequencies around 1,700 Hz (iii) but no responses were observed at 10 kHz (iv). For these experiments, the recording electrodes were carefully shielded from the crystal. The peak-to-peak displacement of the crystal was 1.1 nm.

contralateral auditory nerve is cut, then all responses to the low and intermediate sound frequencies disappear, while those to high frequency tones persist, although their amplitude is about halved (Fig. 4*a*). On the other hand, the amplitude of evoked responses in the midbrain ipsilateral to the cut auditory nerve remain unaffected, except for a slight reduction in the rostral region. Thus, the rostral region of the midbrain receives binocular inputs, in a way analogous to the projection of retinal inputs to the rostral tectum^{16,17}.

Resonance characteristics of the auditory system

One of the peculiar features of the physiology of the auditory system is the enormous variation in its sensitivity to different frequencies. In the frog, there are two major frequency domains, one around 300 Hz and the other around 1,700 Hz, to which the auditory system is especially sensitive⁵, and the neural centre remains unexcited by sound frequencies beyond 3,500 Hz (refs 15, 18). It is of interest to determine whether this differential sensitivity is due to the mechanical properties of the middle ear or to the auditory neurones.

We can demonstrate in several ways that such a frequency response is largely determined by the resonance characteristics of the mouth cavity. For example, if the mouth cavity is open or filled with cotton wool or the Eustachian tubes are blocked, then the characteristic mechanical tuning curve of the eardrum is drastically altered¹².

To demonstrate that the limits of hearing are in part determined by the failure of the eardrum to transmit the incident pressure to the inner ear, we have applied a forced mechanical vibration directly to the eardrum and monitored neural signals from the auditory midbrain. The tip of a fine glass rod fixed to a piezoelectric crystal was gently pressed against the eardrum, and a sinusoidal electric field was then applied to the crystal. Theoretically, the expansion of the crystal is related to the applied electric field across it by the formula

$$dl = e_{31} (l \times V) / z$$

where l and z are the length and thickness, V is the applied voltage, and e_{31} is the piezoelectric coefficient, which, for the crystal we used, was $1.71 \times 10^{-11} \text{ m V}^{-1}$. With this arrangement, we could induce a forced vibration of the eardrum at any known amplitude and frequency. The results of such experiments show that the variation of neural responses observed with airborne frequencies largely disappears if the pressure is directly applied to the eardrum. Moreover, neural responses can be obtained

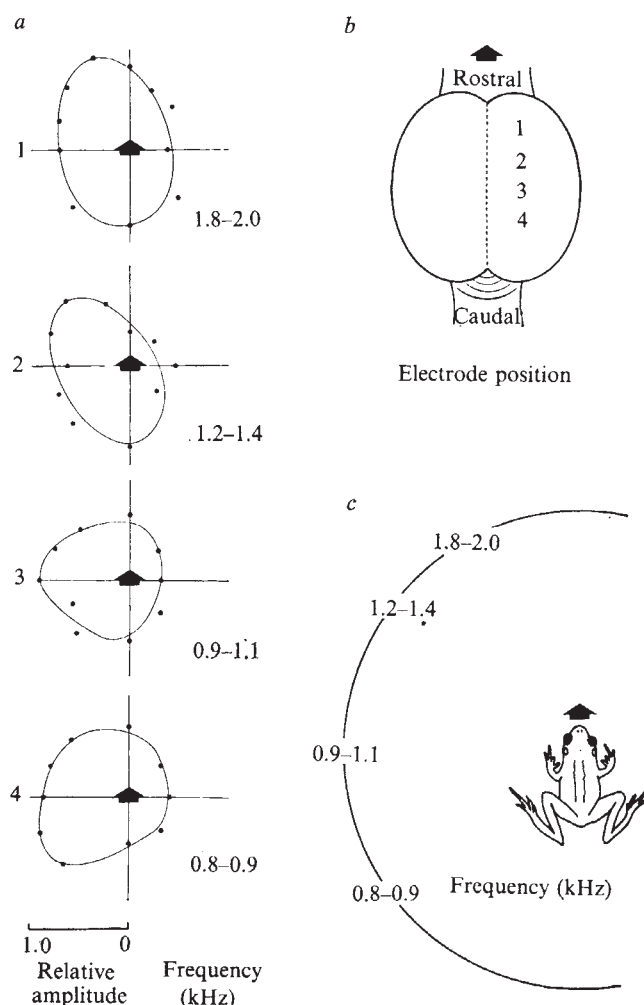


Fig. 5 Variation of the evoked responses with the direction of incident sound. For each of four electrode positions along the anterior-posterior axis of the torus semicircularis, the amplitudes of responses were measured as a function of the direction of incident sound. The relative amplitudes of the measured responses are plotted on polar coordinates (*a*). The distance from the origin represents the relative amplitude, while the position of each dot relative to the origin represents the direction of the sound in relation to the animal. Each of the four figures shown in *a* is obtained from the correspondingly labelled midbrain position shown in *b*. The optimum frequency and direction of the stimulus for the four electrode positions along the anterior-posterior axis of the midbrain are shown in *c*. Note that a given position in the midbrain responds best to a specific frequency range from a certain direction. The distance between the speaker and the animal was 85 cm; the sound pressure used for this experiment was 0.02 Pa.

even when the eardrum is artificially vibrated at 10,000 Hz (Fig. 4b); with an airborne sound at this frequency, the midbrain auditory centre remains unresponsive.

We have also monitored neural responses from the mid-caudal region of the torus semicircularis ipsilateral to the eardrum which was artificially vibrated. The range of frequency response was limited to that observed with airborne stimuli (Fig. 4b). We believe that these responses are the result of transmission of movement from the forced ipsilateral eardrum to the contralateral eardrum through the Eustachian tubes and mouth cavity. The observation that forced vibration above 4,000 Hz does not give a response in the ipsilateral torus semicircularis implies that the movement at these frequencies is not transmitted. It is likely that the frequency range of transmitted movement is limited by the resonance characteristics of the mouth cavity, and this would also explain the normally limited range of frequency sensitivity of the frog's auditory system.

In this context, it is relevant to note that the neural responses to low tones remain relatively unaffected when the eardrum is removed¹⁹. By monitoring neural responses in the midbrain, we have confirmed that the reduction in sensitivity to low tones after cauterisation of the eardrum is small, whereas that for high tones is substantial. For example, the amplitude of neural response elicited by a 150 Hz tone at 0.3 Pa was reduced by 16% when the eardrum was cauterised, whereas the corresponding reduction for 1,750 Hz at 0.03 Pa was 50%. This suggests that sound of low frequency is transmitted to the inner ear by means other than the eardrum.

Map of the auditory space in the midbrain

Whatever the precise mechanisms for directional hearing may prove to be, auditory space is likely to be represented in the midbrain in a way analogous to the map of the visual space on the tectal surface. Several investigators²⁰⁻²² have already noted that in the mammalian tectum the optimal position, relative to an animal's head, for eliciting a response in a particular part of the tectum, is the same for both a sound and a visual object. The following experiments demonstrate that the map of the auditory space in the torus semicircularis is topographically coincident with the representation of the visual space in the tectum.

The frog was placed on a table, the top of which could be freely rotated in the horizontal plane. The distance between the left eardrum and the source of the sound was kept constant (85 cm) throughout the experiment. The sound pressure impinging on the eardrum was continuously monitored with a small omnidirectional pressure transducer suspended just above the eardrum. With this arrangement, we have systematically explored the changes in the amplitude of the evoked potentials as the position of the animal relative to the source of sound was changed.

The amplitude of the evoked responses in the torus semicircularis varies with the direction of the incident sound. If the

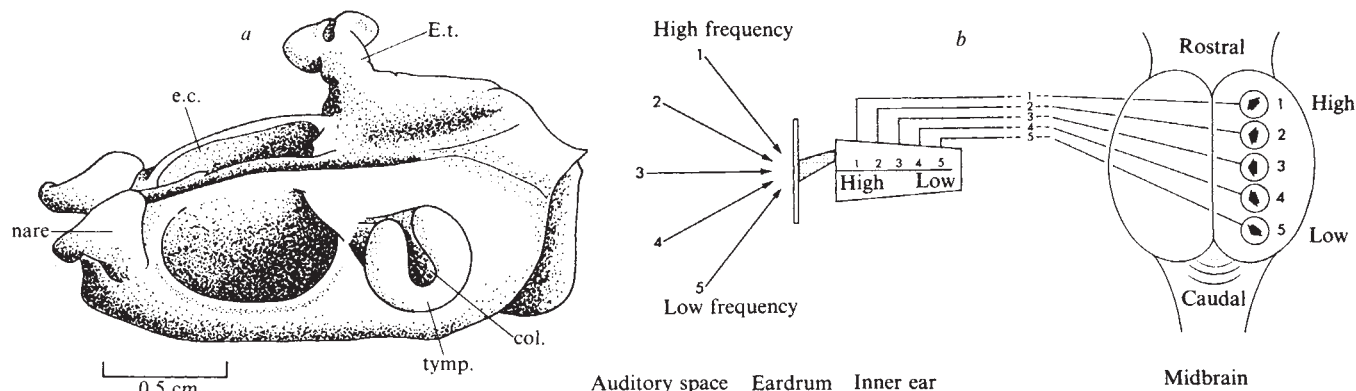
electrode is placed in the right side of the midbrain, sound emanating from the left side of the animal evokes better responses than sound from the right side, although the driving force of the stimulus at the eardrum remains unchanged. The neurones in the auditory midbrain not only respond preferentially to the stimulus in the contralateral field, but they also show directional selectivity within this field (Fig. 5). For example, when the recording electrode is at the rostral region of the midbrain, the best responses are obtained from a high tone originating in the rostral aspect of the contralateral auditory field. On the other hand, if the electrode is placed near the middle of the midbrain, an intermediate tone originating in the auditory field directly facing the eardrum elicits the best responses. Finally, the caudal region of the midbrain is most effectively stimulated by a low tone from the caudal part of the contralateral auditory field. Thus, discrete regions of neural tissue along the anterior-posterior axis of the torus semicircularis act as resonators, each one broadly tuned to a different range of frequencies; and each one responds optimally to the incident sound from a different direction. In this way, a map of auditory space is represented in the midbrain.

Design of the amphibian auditory system

We suggest that the mouth cavity, to which the two ear cavities are connected through Eustachian tubes (Fig. 6a), is acting as a Helmholtz resonator²³. This has several interesting physical consequences. As von Békésy²⁴ first pointed out, it makes possible a reduction in the animal's sensitivity to its own voice. Because the Eustachian tubes of the frog are permanently open, the sound pressure from its own voice will act on both sides of the eardrum simultaneously, thereby minimising the net vibration. On the same principle, the animal's sensitivity to sound from the outside will be greatly diminished if its mouth is open, and this may be one of the reasons why its mouth is always kept tightly shut unless there is good reason to open it. Moreover, this physical arrangement of the ear connected to the mouth cavity may explain why the spectral composition of the mating-call closely matches the sensitivity characteristics of the auditory system^{25,26} (see also Fig. 3); the same resonator is used for producing the mating-call and for modulating the vibration of the eardrum to sound coming from the outside. On this simple principle, selective communication within a species is achieved without any elaborate design in the nervous system.

There are several reasons for supposing that the mechanical properties of the peripheral auditory apparatus form the basis of directional hearing. If this task is performed by the neurones along the auditory pathways, we would expect profuse convergence of the auditory inputs from the two ears. However, our physiological studies (see Fig. 4) and anatomical studies using horseradish peroxidase, have failed to reveal any such convergence in torus semicircularis. Moreover, it is difficult

Fig. 6 Design of the auditory system. *a*, The shape of the mouth and ear cavities is drawn from a space-filling model, which was obtained by injecting dental cement into the mouth cavity. Abbreviations: E.t., Eustachian tube; e.c., indentation resulting from the eyecup; tym., tympanum; col., columella. *b*, A suggested mechanism by which the auditory space is projected on to the midbrain is shown schematically.



to conceive how a temporal difference of 40 μ s can effectively be utilised by neurones, in which any elementary synaptic event must take several ms. On the other hand, if the eardrums are elastically coupled, so that the displacement of one eardrum causes a concomitant displacement of the other, as proposed first by Strother²⁷ for the frog and later by Hill & Boyan²⁸ for the cricket, then the amount of genetic information needed to achieve directional hearing becomes much diminished.

The pattern of connections between the auditory nerve and the midbrain shows a striking similarity to that between the optic nerve and the tectum. In both modalities, one side of the midbrain attends to contralateral sensory space, and the anterior-posterior axis of the external world is represented along the same axis of the midbrain. Moreover, the auditory space directly ahead of the animal is attended by both ears (Fig. 4). This orderly representation of the auditory space is achieved by forming a topologically continuous map of the receptors in the inner ear on to the midbrain. If we assume that the receptors of the amphibian inner ear are tonotopically organised as in higher vertebrates, then the receptors are simply projected onto the midbrain in a correct order, as shown schematically in Fig. 6b. As we show in the following article¹², the eardrum is optimally tuned to a specific tone coming from a region in the auditory space, such that a high tone from the rostral field excites the rostral midbrain and a low tone from the caudal field excites the caudal midbrain. This is an excellent physiological arrangement for directional hearing of the mating-call. The inner ear and the midbrain are performing a separate Fourier analysis of the sound. With the low and intermediate frequency components, the animal can determine on which side its mating partners are located; once the animal orientates towards the source of the sound, it can then utilise the high-

frequency components of the mating-call for guidance. This simple, but ingenious, design has proved to be effective in serving its function, as evidenced by the survival of this species for so many millennia.

We thank Simon J. Caidan for technical assistance. A.G.P. is the recipient of a Nuffield Foundation postdoctoral travelling Fellowship.

Received 20 November 1977; accepted 6 February 1978.

1. Stevens, S. S. & Newman, G. B. *Proc. natn. Acad. Sci. U.S.A.* 20, 593-596 (1934); *Am. J. Psychol.* 48, 297-306 (1936); *Proc. natn. Acad. Sci. U.S.A.* 22, 668-762 (1936).
2. Wallach, H. E., Newman, E. B. & Rosenzweig, M. R. *Am. J. Psychol.* 62, 315-336 (1949).
3. von Hornbostel, E. M. & Wertheimer, M. *S.B. Preuss. Akad. Wiss.* 20, 388-396 (1920).
4. Strother, W. F. *J. aud. Res.* 2, 279-286 (1962).
5. Frishkopf, L. S., Capranica, R. R. & Goldstein, M. H. *Proc. IEEE* 56, 969-980 (1968).
6. Feng, A. S., Gerhardt, H. C. & Capranica, R. R. *J. comp. Physiol.* 107, 241-252 (1976).
7. Harrison, J. M. *Fedn Proc.* 33, 1901-1903 (1974).
8. Masterton, R. B. *Fedn Proc.* 33, 1904-1910 (1974).
9. Starr, A. *Fedn Proc.* 33, 1911-1914 (1974).
10. Moushegian, G. & Rupert, A. L. *Fedn Proc.* 33, 1924-1927 (1974).
11. Erulkar, S. D. *Fedn Proc.* 33, 1928-1932 (1974).
12. Chung, S. H., Pettigrew, A. G. & Anson, M. *Nature* 272, 142-147 (1978).
13. Gesteland, R. C., Howland, B., Lettvin, J. J. & Pitts, W. H. *Proc. Inst. Radio. Engrs* 45, 1856-1862 (1959).
14. Freeman, J. A. *Nature* 269, 218-222 (1977).
15. Potter, D. J. *Neurophysiol.* 28, 1132-1154 (1965); 28, 1155-1184 (1965).
16. Gaze, R. M. *J. Physiol., Lond.* 143, 20P (1958).
17. Gaze, R. M. & Jacobson, M. Q. *Jl exp. Physiol.* 47, 273-280 (1962).
18. Frishkopf, L. S. & Goldstein, M. H. *J. acoust. Soc. Am.* 35, 1219-1228 (1963).
19. Lombard, R. E. & Straughan, I. R. *J. exp. Biol.* 61, 71-93 (1974).
20. Gordon, B. *J. Neurophysiol.* 36, 157-178 (1973).
21. Dräger, U. C. & Hubel, D. H. *Nature* 253, 203-204 (1975).
22. Chalupa, L. M. & Rhoades, R. W. *J. Physiol., Lond.* 270, 595-626 (1977).
23. Helmholtz, H. J. *rein. angew. Math.* 57, 1-72 (1860).
24. von Békésy, G. *Experiments on Hearing* (McGraw-Hill, New York, 1958).
25. Capranica, R. R. *Res. Monogr.* 33 (MIT Press, Cambridge, Massachusetts, 1965).
26. Loftus-Hills, J. J. & Johnstone, B. M. *J. acoust. Soc. Amer.* 47, 1131-1138 (1969).
27. Strother, W. F. *J. comp. Physiol. Psychol.* 52, 157-162 (1959).
28. Hill, K. G. & Boyan, G. S. *Nature* 262, 390-391 (1976).

Dynamics of the amphibian middle ear

Shin-Ho Chung, Alan Pettigrew & Michael Anson

National Institute for Medical Research, London NW7, UK

Using laser speckle interferometry, we show that the directional information of acoustic signals is encoded in the motion of the eardrum. Moreover, the frequency sensitivity of the auditory system is determined by the mechanical properties of the middle ear, which acts like a damped resonator.

ON the basis of our electrophysiological study of the auditory system of *Rana temporaria*¹, we have made several predictions about the movement of the eardrum in response to acoustic stimuli. First, the mouth cavity to which the two ears are connected through Eustachian tubes acts as a Helmholtz resonator², and determines the frequency responses of the peripheral auditory system. Second, the pressure generated by the displacement of one eardrum will be transmitted through this resonator to the other eardrum, the net vibration of which may thereby be reduced. Finally, the amplitude of the eardrum's movement will critically depend on the location of the sound source relative to the animal.

We have tested these predictions using laser speckle interferometry. Here we report the results of our study, which verify all our predictions. In addition, we have analysed transient motions of the eardrum to derive all the numerical constants concerning the mechanical properties of the middle ear.

Laser speckle interferometry

Several previous measurements on the invertebrate³ and vertebrate⁴⁻⁹ middle ear have shown that the displacement of the

eardrum at a sound pressure of 1 Pa is of the order of less than 20 nm. To monitor such a small movement without distorting the mechanical properties of the vibrating parts, we have developed a simple technique of coherent speckle interferometry. A polarised He-Ne laser (2 mW, $\lambda = 632.8$ nm) was directed at the eardrum, and scattered radiation was monitored with a photodiode. A beam splitter placed near the output mirror of the laser directed about 10% of the laser beam to a second photodiode. The photocurrents from each of the detectors were amplified with current-to-voltage amplifiers¹⁰, and the amplified output from the signal detector was divided by that from the second detector with a specially designed voltage divider. This ratioing operation reduces the effective amplitude fluctuations of the laser by more than 100-fold¹¹. The divided signal was then fed into a spectrum analyser (Marconi TF 2370), or a signal averager (Biomac 1000), or a phase-sensitive detector (Brookdeal 401A). A block diagram of our experimental arrangement is shown in Fig. 1a.

If coherent light is scattered by an irregular surface, such as the eardrum, then the scattered field forms dark and light fringes, known as the speckle pattern¹². The speckle field scattered from an oscillating surface is modulated in phase and, when mixed with an unperturbed reference field, will give an amplitude fluctuation at the photodetector. In practice, it is not necessary to provide the reference beam deliberately; for example, the amount of unmodulated light needed to generate 0.1% fluctuation of the detected signal is about 6×10^{-5} of the modulated light, and this amount is provided in our experiment by stray light. If the scattering surface undergoes sinusoidal vibrations with amplitude a_0 at angular frequency ω_a , then the

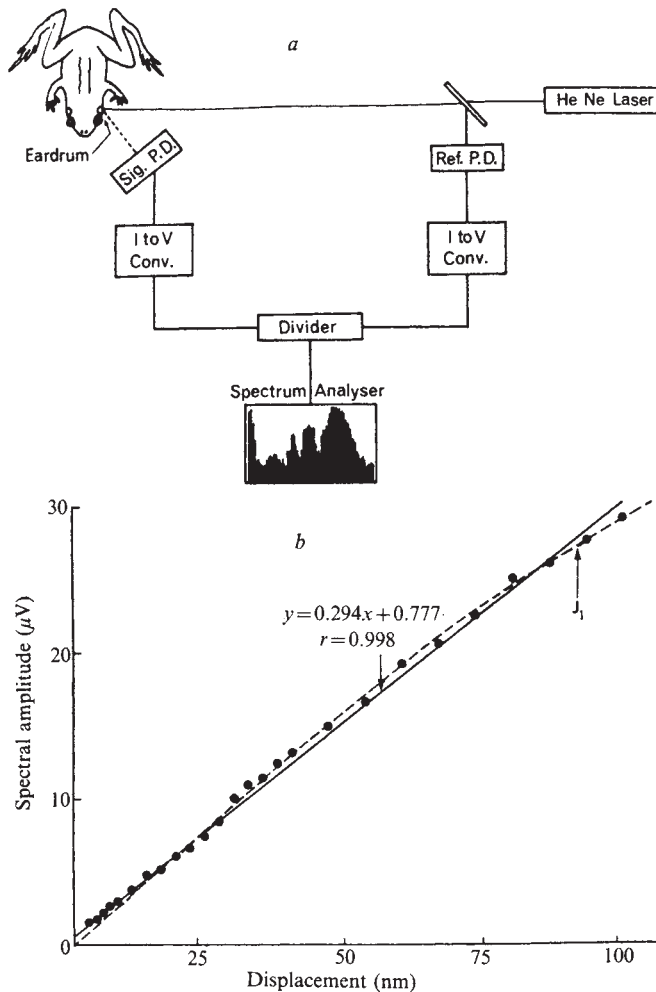


Fig. 1 Laser speckle interferometry. *a*, The experimental arrangement is shown schematically. A polarised He-Ne laser (Spectra-Physics, Model 133) was placed at 65 cm from the animal, while the photodiode monitoring scattered radiation was positioned 2 cm from the eardrum. The vertical angle between scattered and incident beams was about 30°. The output and input impedances of the various electronic instruments were matched with buffer amplifiers. *b*, The spectral amplitude is plotted against peak-to-peak displacement of a piezoelectric crystal, whose expansion coefficient was $1.71 \times 10^{-11} \text{ m V}^{-1}$. Frequency of oscillation was 1,000 Hz. In this and the following figures, each dot represents an average of 10 measurements, and the solid line is a least-squares fit of the experimental data. The broken line is the computed value of the Bessel function $J_1(2\pi x/\lambda)$.

with photocurrent at the signal detector with responsivity σ is:

$$I = \sigma [\langle \mathbf{E}_s \cdot \mathbf{E}_s^* \rangle + \langle \mathbf{E}_r \cdot \mathbf{E}_r^* \rangle + 2 | \mathbf{E}_s | \cdot | \mathbf{E}_r | \cos ((4\pi a_0/\lambda) \sin \omega_a t + \phi)]$$

where $\mathbf{E}_s$ and $\mathbf{E}_r$ are, respectively, total electric fields at the detector due to the scattered and reference signals, stars denote complex conjugates, and ϕ is an arbitrary phase difference between the signal and reference. The third term in the equation may be expanded as a series of Bessel functions and we may select the component at the fundamental frequency ω_a such that:

$$I(\omega_a) = \sigma [| \mathbf{E}_s | \cdot | \mathbf{E}_r | J_1(4\pi a_0/\lambda) \sin \omega_a t \cdot \sin \phi]$$

where J_1 is a Bessel function of the first kind and first order with argument $4\pi a_0/\lambda$. In words, the amplitude of the photocurrent at the detector will be maximal if the scattering surface undergoes a sinusoidal motion with peak-to-peak displacement

of 185 nm, and minimum with peak-to-peak displacement of 386 nm.

To ascertain that our optical method of measuring displacement conforms with the theoretical predictions given above, we monitored the displacement of a piezoelectric crystal (PZT 5A, Vernitron) which could be vibrated at a known amplitude and frequency (see Pettigrew *et al.*¹). As the applied sinusoidal voltage across the piezoelectric crystal increased, the amplitude of the detected signal at the vibrating frequency first increased and then decreased. The peak amplitude of the response was obtained when the peak-to-peak displacement was about 210 nm; the response diminished when the displacement of the piezoelectric crystal was further increased. From several measurements, we have ascertained that the response was approximately linear over displacement up to 100 nm (Fig. 1*b*).

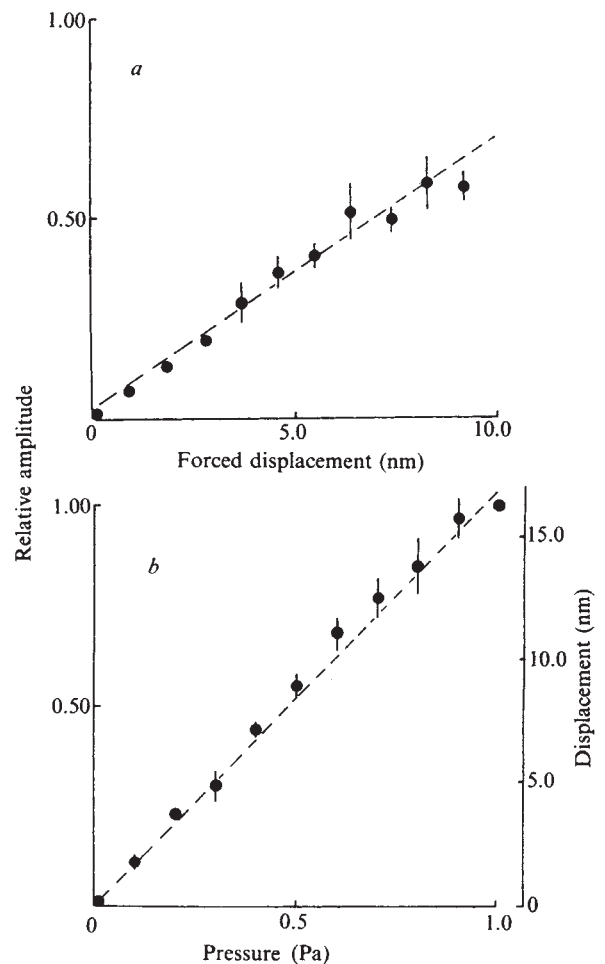


Fig. 2 Displacement of the eardrum. *a*, The spectral amplitude was measured while the eardrum was forced to vibrate at known amplitudes. To induce the movement of the eardrum at a known amplitude, the tip of a fine glass rod attached to a piezoelectric crystal was pressed gently against the middle of the eardrum, and a sinusoidal electric field was applied across the crystal. When the tip of the glass rod was vibrating but not in contact with the eardrum, no responses were observed. *b*, The spectral amplitudes at various sound pressures were then measured, and the averaged values over 3 different experiments are plotted. The frequency of the stimulus was 1,600 Hz throughout. From the slopes of the graphs in *a* and *b*, we calculated the absolute displacement as a function of sound pressure, as shown on the right hand side ordinate in *b*. We note here that sound pressures as well as displacement are expressed as root-mean-squares. If sound pressure is measured as peak-to-peak, the ordinate in *b* represents peak-to-peak displacement. For (*a*), $y = 0.07x + 0.03$; for (*b*), $y = 1.04x + 0.02$.

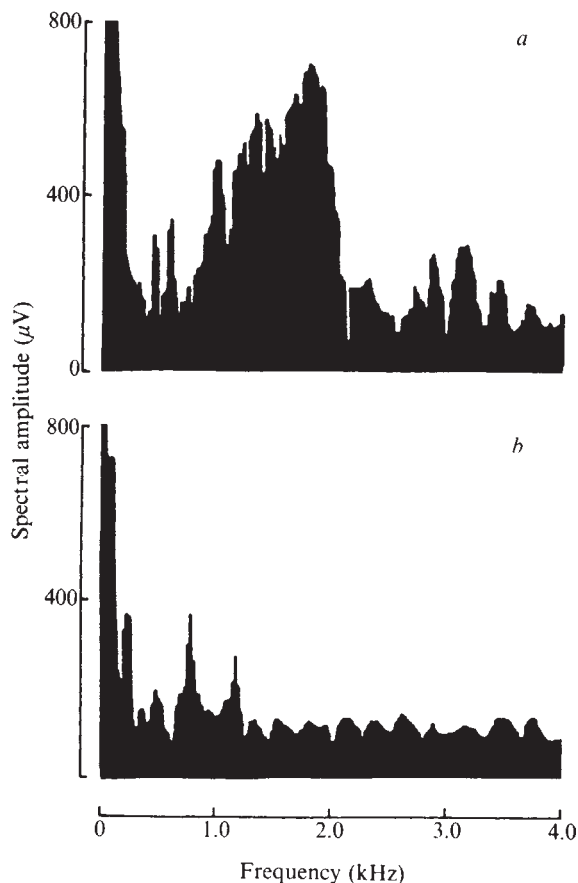


Fig. 3 Resonance characteristics of the middle ear. *a*, The spectral amplitude was recorded while the frequency of sound at a constant pressure (0.5 Pa) was modulated from 0 to 5,000 Hz. The bandwidth of the spectrum analyser was set at 50 Hz. The sharp peak at the frequency near zero is an artefact of the spectrum analyser. *b*, When the same measurement was made after the mouth was forced open, the eardrum virtually ceased to vibrate.

Resonance characteristics of the middle ear

We first measured the displacement of the eardrum in response to different forces impinging on it. With a fine glass rod attached to a piezoelectric crystal, the eardrum was forced to

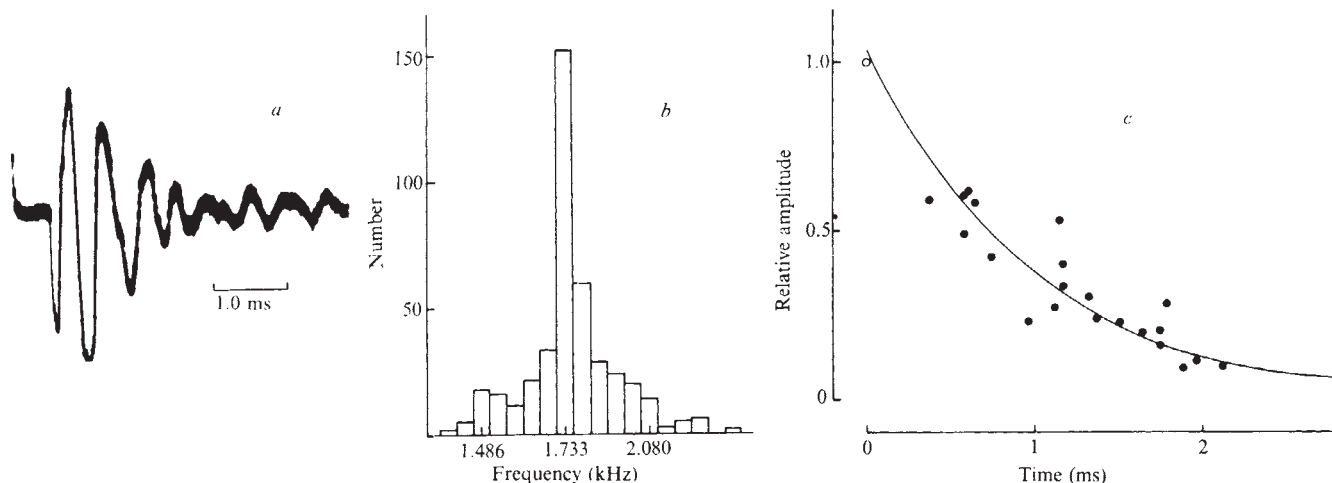
vibrate at 1,600 Hz, while the a.c. component of the modulated light was analysed from 0 to 5,000 Hz with bandwidths of 50 Hz. Against a flat background, a prominent peak appeared in the spectrum analyser at the vibrating frequency, and the amplitude of this response increased with increasing displacement of the eardrum. As shown in Fig. 2*a*, the relationship between the amplitude of response and the amplitude of displacement was approximately linear. The glass rod was then removed, and a similar strength-response curve was constructed by vibrating the eardrum with sound at the same frequency. This relationship too could be reasonably well fitted with a linear function for sound pressure up to 1 Pa (Fig. 2*b*). From these two linear functions, we have calculated the absolute displacement of the eardrum as a function of sound pressure, as shown in Fig. 2. For example, a sinusoidal pressure of 1 Pa at 1,600 Hz displaces the eardrum by 16.1 nm; the maximum discrepancy of this estimate from different experiments was 19%. Similar estimates have previously been made on other invertebrate³ and vertebrate⁴⁻⁹ species using different measuring techniques. We therefore conclude that, within the physiological range of sound intensity, the eardrum obeys Hooke's law.

The amplitude of displacement for a fixed sound pressure depends on the frequency of the driving force. In Fig. 3*a*, the frequency of sound at a constant pressure (0.5 Pa) was modulated from 0 to 5,000 Hz, and the tuning of the spectrum analyser was locked to this signal. Extrapolating from the measurements shown in Fig. 2, the amplitude of the analysed signal at each frequency represents the relative displacement of the eardrum in response to sound at that frequency; the peak amplitude of vibrations occurs at about 2,000 Hz, while the eardrum remains relatively unresponsive to frequencies below 1,000 Hz and above 4,000 Hz. If similar measurements are made either with the mouth cavity forced open or filled with cotton wool, or with the Eustachian tubes blocked with dental wax, then the sensitivity of the eardrum declines drastically and the variation in the amplitude of its response with frequency becomes much diminished (Fig. 3*b*). We therefore conclude that the mouth cavity acts as a resonator, whose characteristics largely determine the frequency selectivity of the middle ear.

Damped oscillations of transients

If the movement of the eardrum obeys Hooke's law (compare Fig. 2), then it may be possible to treat its motion as that of a damped resonator. Our strategy was to estimate, from the analysis of transient motions, the numerical constants of the differential equation describing a harmonic oscillation with damping.

Fig. 4 Damped oscillations of the eardrum. *a*, In response to the sound of an electric spark (see text for details), the eardrum executed several rapid oscillations. The pressure generated by the electric spark initiated the damped oscillations of the eardrum. *b*, To estimate the natural frequency of the middle ear, the period of each oscillation was measured and tabulated. Note that the predominant frequency is 1,733 Hz. *c*, To estimate the friction coefficient of the middle ear, the amplitude of each oscillation, normalised to an initial amplitude of unity, was measured and plotted against time (in ms). The exponential function was fitted through the points by the least-square method. For (*c*), $y = 1.03 \exp(-1.04t)$.



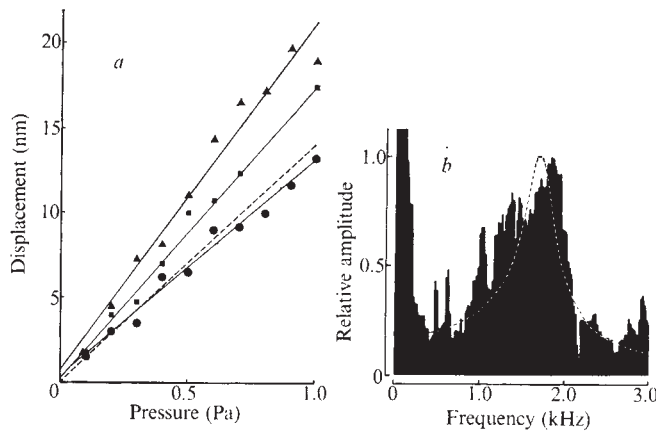


Fig. 5 Dynamics of the middle ear. *a*, From the numerical estimates obtained from the analysis of damped oscillations, a strength-response curve was predicted (broken line). The observed values from 3 separate experiments are fitted by solid lines. The mass of the vibrating parts is that of the eardrum and the columella. $\kappa = 3.59 \times 10^3$ (kg s⁻²); $A = 9.62 \times 10^{-6}$ (m²); $Q = 5.26$. *b*, The predicted resonance characteristics of the middle ear (broken line) is superimposed on the measured resonance curve. For details, see Fig. 3. $m = 3 \times 10^{-5}$ (kg); $\gamma = 2.08 \times 10^3$ (s⁻¹).

Assuming that the dynamics of the middle ear can be represented by this equation of motion, we have predicted from the numerical constants the behaviour of the eardrum under various conditions of driving force and frequency, and then compared theoretical predictions with direct measurements. Thus, the analysis of transients permits us to verify our assumption and, at the same time, allows us to check the reliability of our direct measurements with an independent method.

To derive these constants, the ear was stimulated by the sound of an electric spark. A 0.2 μ F capacitor was charged to 22,000 V in 4 s and then allowed to discharge through a 7 mm spark gap in 10 μ s. In response to the instantaneous pressure of the sound, the eardrum executes several oscillatory motions, with the amplitude of each successive oscillation declining rapidly (Fig. 4a). The solution of the differential equation for the damped vibration is of the form:

$$X = \exp(-\gamma t/2) \cos \omega_r t$$

where the amplitude of initial oscillation is normalised to unity, and frequency of damped oscillation ω_r is related to the resonance frequency ω_0 by $\omega_r^2 = \omega_0^2 - \gamma^2/4$. Thus, the damped transient motion is an oscillation of frequency close to the resonance frequency, and the amplitude of this motion diminishes as $\exp(-\gamma t/2)$. We have estimated ω_r and the friction coefficient γ from photographic records of the transient motions. The period, $2\pi/\omega_r$, of each oscillation was plotted as a histogram (Fig. 4b), and the amplitude of each successive oscillation was plotted against time (Fig. 4c). Our estimated values for the resonance frequency and friction coefficient are, respectively, 1,741 Hz and 2.08×10^3 s⁻¹.

Once γ and ω_0 are known, we can derive the stiffness κ of the system, since $\omega_0^2 = \kappa/m$, where m is the mass of the vibrating parts. The value κ is estimated to be 3.59×10^3 kg s⁻². These numerical constants provide a complete prediction of the motion of the eardrum such that the displacement of the membrane under the driving force F at frequency ω close to the resonance frequency is:

$$X = (F/\kappa) Q$$

where Q is the resonance enhancement factor. Moreover, we can predict the resonance curve of the system with the formula:

$$X^2 \propto 1/m^2[(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2]$$

It also follows that the displacement X is lagging the force F by:

$$\theta = \tan^{-1} \left(\frac{-\gamma\omega}{\omega_0^2 - \omega^2} \right)$$

and careful analysis of this phase shift, we believe, will enable us to formulate an accurate mathematical model of the amphibian auditory system.

As shown in Fig. 5, there is a good agreement between our prediction and actual measurements of the intensity-response curve and resonance curve. The observed phase shift, not shown here, also conformed with the theoretical prediction; the phase angle between the driving force and the motion of the eardrum gradually shifted from 0 to $-\pi$ as a function of frequency, reaching the midpoint at resonance. This close agreement suggests that, in the first approximation, the dynamics of the middle ear can conveniently be treated as a damped resonator. Although the overall behaviour of the eardrum obeys the simple equation of motion, there are several peculiar features of vibration which cannot be accounted for without introducing an additional term to the equation. For example, the motion of the eardrum undergoes a periodic fluctuation in amplitude and phase in response to a steady tone. The physical basis for this interference, along with a mathematical analysis, will be given elsewhere.

We have also calculated the resonance frequency of the mouth cavity and Eustachian tubes, assuming that they act like a Helmholtz resonator. Helmholtz² deduced that a resonator with the volume S will resonate at a frequency f according to the theoretical formula:

$$f = \frac{c A^{\frac{1}{2}}}{2^{\frac{1}{2}} \pi^{\frac{1}{2}} S^{\frac{1}{2}}}$$

where c is the speed of sound in dry air, and A is the cross sectional area of Eustachian tube. He also provides an empirical formula, which assumes slightly different numerical constants from those in the theoretical formula. The resonance frequency of the frog's middle ear predicted from the theoretical formula is 1,886 Hz, whereas that calculated from the empirical formula is 1,666 Hz. These values are within 10% of our measured frequency (compare Fig. 4b).

The two eardrums act as coupled pendula

If the two eardrums are mechanically coupled so that force impinging on one is transmitted to the other through the closed intracranial air space, then the motion of the eardrum will show directional sensitivity. Thus, the eardrum contralateral to the incident sound will be displaced outward at the time when the sound induces inward movement of this eardrum. Because pressure is acting on both sides of the membrane simultaneously, its net vibration will thereby be diminished, as proposed previously^{13,14}. Here we provide direct evidence that the eardrums are elastically coupled.

The vibration of one eardrum was monitored while the other was forced to oscillate with the piezoelectric crystal. As the amplitude of the displacement on one eardrum increased above 4 nm, the amplitude of the transmitted displacement also increased. In this way, the two membranes oscillated in phase with the driving frequency of the crystal. We have made a number of measurements to determine the efficiency of this elastic coupling. First, we measured the relative amplitude of the transmitted movement as a function of the applied movement on the other eardrum. The frequency of this applied vibration was kept constant throughout at 1,600 Hz. Then, the intensity-response curve for sound of the same frequency was constructed, so that we could infer the absolute amplitude of the transmitted movement. In two such experiments shown in

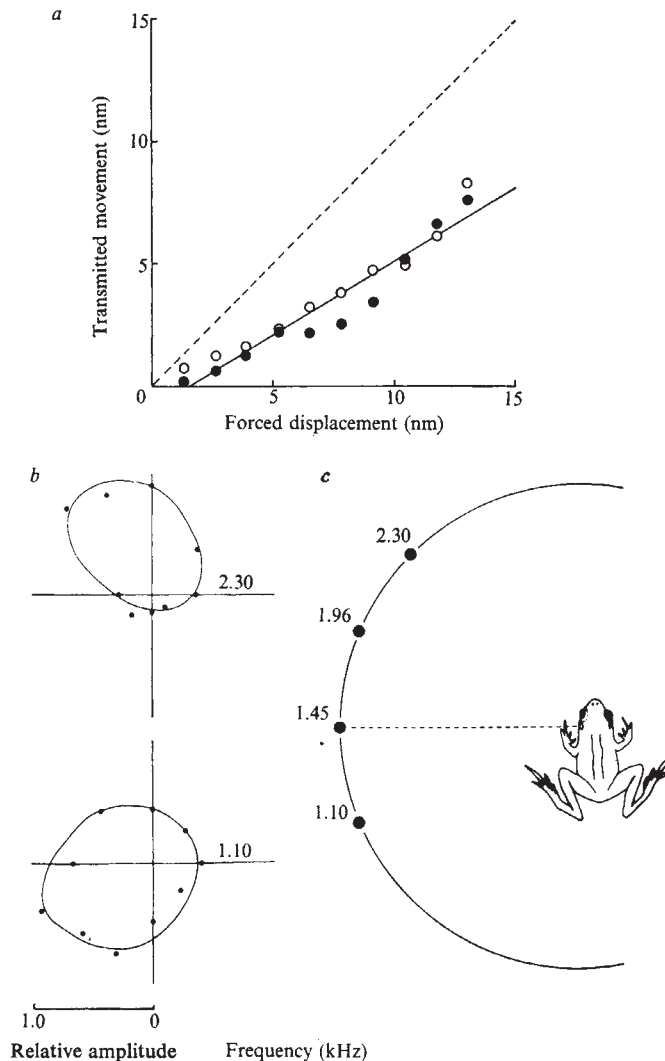


Fig. 6 Two eardrums acting as coupled pendula. *a*, One eardrum was forced to oscillate at 1,600 Hz with a piezoelectric crystal, while we monitored the amplitude of the transmitted movement of the other eardrum. The amplitude of transmitted movement is plotted as a function of the amplitude of forced movement. To ascertain the absolute amplitude of transmitted movement, we used the procedure detailed in Fig. 2. The results shown here are from two separate experiments. *b* and *c*, The displacement of the eardrum was measured at a number of frequencies while the animal was rotated along the horizontal plane in relation to a fixed sound source. The amplitude of the eardrum's displacement is plotted as a function of direction of incident sound in the polar co-ordinate in *b*. The optimum directions for 4 different frequencies are shown in *c*.

Fig. 6a, the coefficient of the coupling was 0.5; for example, a forced oscillation of one eardrum by 10 nm induced a symmetrical oscillation in the other eardrum by 5 nm.

In the following experiments, we show that the amplitude of vibration of the eardrum is critically dependent on the direction of incident sound. The animal was mounted on an elevated platform on a table which could freely rotate in the horizontal plane, so that one eardrum occupied the centre of the axis of rotation. To minimise the acoustic shadow created by the optical instruments mounted on the same table, the laser light was directed to the eardrum with a pair of first-surface mirrors. With this arrangement, we could alter the direction of incident sound relative to the animal's body axis by rotating the table. The magnitude of the eardrum's response varied systematically when the direction of incident sound was changed, as shown in Fig. 6b and c. For all frequencies tested, the sound originating from the auditory space ipsilateral to the eardrum always caused more effective vibration than that from

the contralateral field. Even within this ipsilateral field, each frequency of sound had an optimum position for eliciting the maximum vibration. For example, for the high tone of 2,300 Hz, the best response was obtained when the incident sound originated from the rostral auditory field, whereas the optimal position for intermediate tones was more caudal. We find in general that the most effective vibration of the eardrum is obtained when the pressure wave acting on the outside of the membrane leads the pressure wave acting on the inside of the same membrane by 90°. In symbols,

$$D + d \cos \theta = \lambda/4$$

where D and d are, respectively, the internal and external distances between the eardrums, and θ is the angle of the sound source relative to the line intersecting the two ears. Experimental evidence available at the moment does not allow us to discriminate between several possible physical interpretations of this empirical relationship. The asymmetrical responses between the rostral and caudal auditory fields are probably due to a slight inclination of the eardrums towards the rostral field. In this way, the motion of the eardrum faithfully reflects the direction of the incoming sound.

The two auditory systems

Our study clearly demonstrates that the mechanical properties of peripheral auditory apparatus determine several behavioural and electrophysiological characteristics of the amphibian auditory system. Among these are the limits of hearing, the variation in sensitivity with frequency, and the perception of directional hearing. For mathematical analysis, we have treated the dynamics of the middle ear as a damped resonator, whose resonance frequency is governed in part by the mouth and ear cavities. The numerical constants which have been extracted from the data will need revision as more accurate measurements are made. Moreover, we have treated the mouth cavity as a resonator with one opening, and considered that the restoring force is linear with displacement. We emphasise that the analysis without these simplifying assumptions becomes more complicated than that which we have shown here, but describes the behaviour of the eardrum under certain conditions more adequately.

There is, however, one major difference between the sensitivity of the eardrum and that of the auditory system as a whole. Whereas the eardrum vibrates to the frequency range between 1,000 Hz and 3,000 Hz, the neurones in the midbrain show special sensitivity to the low tones as well. This implies that the pressure resulting from the low frequency sound is transmitted to the inner ear by means other than the motion of the eardrum. Several investigators¹⁵⁻¹⁸ have postulated that the operculus which is vibrated by the active contraction of muscles, is responsible for transducing low frequency sound into the movement of the oval window. Our results, in conjunction with other related studies¹⁸, are consistent with this view. This separate arrangement is necessitated by the physical principle that the amplitude of vibration is inversely proportional to the square of the driving frequency. If we approximate the shape of the frog as a sphere, we can make use of Lichte's formula¹⁹ to estimate the effective displacement of the animal as a function of the driving frequency. Similarly, we can also calculate the displacement of the skin as a function of the driving frequency. Our calculations revealed that only for frequencies below 800 Hz is the magnitude of this displacement substantial enough to be detected by the vestibular and somatosensory systems, and above 1,000 Hz the direct displacement by the impinging force becomes rapidly ineffective. It is for the detection of frequencies above this value that the peripheral auditory apparatus of the frog seems to be specifically designed.

We thank Simon J. Caidan for technical assistance. A. G. Pettigrew is the recipient of a post-doctoral travelling Fellowship from the Nuffield Foundation.

Received 2 November 1977; accepted 6 February 1978.

1. Pettigrew, A., Chung, S. H. & Anson, M. *Nature* **272**, 138–142 (1978).
2. Helmholtz, H. *J. reit. angew. Math.* **57**, 1–72 (1860).
3. Dragsten, P. R., Webb, W. W., Paton, J. A. & Capranica, R. R. *Science* **185**, 55–57 (1974); *J. acoust. Soc. Amer.* **60**, 665–671 (1976).
4. Möller, A. R. *J. acoust. Soc. Amer.* **35**, 1526–1534 (1963).
5. Hillman, O., Schechter, H. & Rubinstein, M. *Rev. mod. Phys.* **36**, 360 (1960).
6. Gilad, P. *et al. J. acoust. Soc. Amer.* **41**, 1232–1236 (1967).
7. Tonndorf, J. & Khanna, S. M. *J. acoust. Soc. Amer.* **44**, 1546–1554 (1968).
8. Johnstone, B. M., Saunders, J. C. & Johnstone, J. R. *Nature* **227**, 625–626 (1970).

9. Høgmøen, K. & Løkkberg, O. J. *Appl. Opt.* **16**, 1869–1875 (1977).
10. Hamstra, R. H. & Wendland, P. *Appl. Opt.* **11**, 1539–1547 (1972).
11. Anson, M. & Bayley, P. *Rev. sci. Instrum.* **47**, 370–373 (1976).
12. Dainty, J. C. (ed.) *Topics in Applied Physics* **9** (Springer-Verlag: Berlin Heidelberg, New York, 1975).
13. Strother, W. F. *J. comp. Physiol. Psychol.* **52**, 157–162 (1959).
14. Hill, K. G. & Boyan, G. S. *Nature* **262**, 390–391 (1976).
15. Kingsbury, B. G. & Reed, H. R. *J. Morph.* **20**, 549–628 (1909).
16. Eiselt, J. *Arch. Naturgesch.* **10**, 179–230 (1941).
17. Baker, M. C. *Copeia*, **3**, 613–616 (1969).
18. Lombard, R. E. & Straughan, I. R. *J. exp. Biol.* **61**, 61–93 (1974).
19. Lichte, N. *Physik. Zeitschr.* **19**, 17–20 (1918).

letters to nature

Re-evaluating Bode's law of planetary magnetism

THE popularity of the so-called magnetic Bode's law for the magnetic dipole moments of the planets^{1–4} persists despite the lack of any physical justification for such a law. The magnetic Bode's law takes the form of a plot of the log of the magnetic moment of a planet against the log of its angular momentum. On this plot, the planetary moments lie roughly along a straight line. But the dimensions of the planets vary widely and both the magnetic moment and angular momentum depend on high powers of these dimensions, so that the apparent correlation of the two quantities should not be surprising. On the positive side, this display illustrates the relative strengths of the magnetic moments of the planets without invoking a particular physical mechanism for the source of those fields, and thus has the advantage of being 'noncontroversial'. We shall here bring up to date the magnetic Bode's law based on recent analysis of the moments of the terrestrial planets.

Figure 1 shows the new Bode's law. The best known moments are those of the Earth ($8 \times 10^{25} \text{ G cm}^3$) and of Jupiter ($1.5 \times 10^{30} \text{ G cm}^3$) (ref. 5). It is natural to join these by a straight line. However, it is not obvious whether we should use the angular momentum of the Earth, or of the Earth–Moon system for plotting the terrestrial value⁴. For example, if we took a precessional source of energy for the terrestrial dynamo^{6,7} then the angular momentum of the Earth–Moon system might be the proper quantity. We have plotted both possibilities. The error bars on the Earth's moment denote the standard deviation of the palaeomagnetic dipole field over the last 8 Myr (ref. 8).

The magnetic moment of Mercury has recently been revised to $2.4 \times 10^{22} \text{ G cm}^3$ (refs 9,10). It lies just below the dashed extrapolation. The Venus moment, however, has just been revised upwards. Estimates range over $3\text{--}6.5 \times 10^{22} \text{ G cm}^3$ (refs 11–13). The moment lies close to the extrapolated solid line. There are two estimates for Mars: the upper one of $2.5 \times 10^{22} \text{ G cm}^3$ is the value originally reported by Dolginov *et al.*¹⁴; the lower value is an upper limit based on critiques^{15,16} of the signatures used by Dolginov *et al.* to identify the martian magnetopause. But there is no conclusive evidence that any of the Mars orbiters has ever encountered a region of intrinsic martian magnetic field. Finally the moment of Saturn as inferred from radio noise observations is $2.2 \times 10^{29} \text{ G cm}^3$ (ref. 17).

As noted above, we attach no physical significance to the fact that the majority of the magnetic moments of the major bodies of the Solar System cluster about the two lines in Fig. 1 or along straight lines in similar scaling laws^{7,18}. In fact, the low value for the martian moment relative to Earth and Mercury suggests that no such scaling law, which ignores the state of the planetary

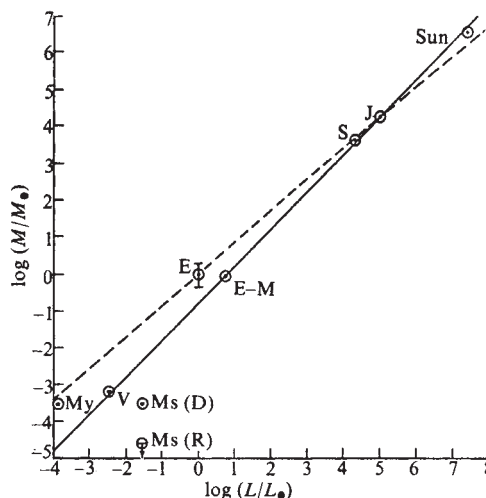


Fig. 1 Bode's law of planetary magnetism. Magnetic moment normalised to terrestrial moment is plotted against angular momentum normalised to terrestrial angular momentum. My, Mercury; V, Venus; Ms, Mars; E, Earth; E-M, Earth–Moon system; J, Jupiter; and S, Saturn. The two estimates for Mars correspond to the original report by Dolginov *et al.*¹³ (D) and my recent re-evaluation¹⁴ (R).

interior, can be made to fit the data. The two major parameters in such laws are spin rate and planetary radius. Mars has a spin rate comparable to the Earth's and a radius intermediate between that of Mercury and that of the Earth. Similarly, even though it is quite likely that Saturn, like Jupiter, has a planetary dynamo, it is not a foregone conclusion that Uranus and Neptune also have. Finally, it is likely that the expected smaller 'conducting core' of Saturn relative to Jupiter¹⁹ results in a greater predominance of the dipole moment relative to the higher moments than on Jupiter. In fact, the relative contributions of the multipole moments of Saturn may be very similar to the terrestrial ratios²⁰.

C. T. RUSSELL

*Institute of Geophysics and Planetary Physics,
University of California,
Los Angeles, California 90024*

Received 20 October; accepted 30 December 1977.

1. Kennel, C. F. *Space Sci. Rev.* **14**, 511–533 (1973).
2. Hill, T. W. & Michel, F. C. *Rev. geophys. Space Phys.* **13**, 967–974 (1975).
3. Dessler, A. J. in *Solar Wind Interactions with the Planets Mercury, Venus, and Mars*, 159–166 (NASA SP-397, Washington, D.C., 1976).
4. Siscoe, G. L. in *Solar System Plasma Physics—A Twentieth Anniversary Review* (eds Kennel, C. F., Lanzetta, L. J. & Parker, E. N.) (in the press).
5. Smith, E. J. *et al. J. geophys. Res.* **79**, 3501–3513 (1974).
6. Malkus, W. V. R. *J. geophys. Res.* **68**, 2871–2886 (1963).
7. Dolginov, Sh. Sh. in *Solar-wind interaction with the Planets Mercury, Venus, and Mars* 167–170 (NASA SP-397, Washington, D.C., 1976).

8. Kono, M. *Earth planet. Sci. Lett.* **11**, 10 (1971).
9. Whang, Y. C. *J. geophys. Res.* **82**, 1024-1030 (1977).
10. Jackson, D. J. & Beard, D. B. *J. geophys. Res.* **82**, 2828-2836 (1977).
11. Russell, C. T. *Geophys. Res. Lett.* **3**, 125-129 (1976).
12. Russell, C. T. *Geophys. Res. Lett.* **3**, 589-592 (1976).
13. Dolginov, Sh. Sh., Zhuzgov, L. N., Sharova, V. A., Buzin, V. B. & Yeroshenko, Ye. G. *The Magnetosphere of Venus* (in Russian) preprint No. 19(193), 65 (IZMIRAN, Moscow, 1977).
14. Dolginov, Sh. Sh., Yeroshenko, Ye. G. & Zhuzgov, L. N. *J. geophys. Res.* **81**, 3353 (1976).
15. Russell, C. T. *Geophys. Res. Lett.* **5**, 81 (1978).
16. Russell, C. T. *Geophys. Res. Lett.* **5**, 85 (1978).
17. Brown, L. W. *Astrophys. J. Lett.* **198**, L89-L92 (1975).
18. Freeman, J. W. in *Proc. NATO Advanced Study Institute on Origin of the Solar System* (in the press).
19. Slattery, W. L. *Icarus* **32**, 58-72 (1977).
20. Elphic, R. C. & Russell, C. T. *Geophys. Res. Lett.* (in the press).

Uranus and the shape of elliptical rings

AN interesting question raised by the discovery¹⁻³ of rings around Uranus concerns the ellipticity of the ϵ ring, the outermost of the five rings detected. When the star SAO158687 passed behind the Uranus system (with a closest approach to the centre of the planet of about 25,000 km) its light was occulted twice by the ϵ ring. The first part of the ring to occult was about 100 km wide and 51,700 km from the planet's centre; the second part was about 40 km wide and 51,050 km from the centre¹ (J. L. Elliot, personal communication). These two points were separated in angle by about 120°. The average transmission of starlight during these events, 35% and 10%, respectively, lead to optical depths that are, within experimental error, inversely proportional to the two widths (J. L. Elliot, personal communication), strongly suggesting that there is the same amount of material in the rings at the two points sampled. Here we account for the variable width of the ring by differences in the orbital eccentricities of the individual particles composing the ring.

A narrow elliptical ring must be composed of particles in elliptical orbits with their major axes in nearly perfect alignment. As the alignment becomes less perfect, the ring width fattens and becomes less elliptical, until complete randomisation of axis directions results in a circular ring.

Partial misalignment of orbital axes could cause the different ring widths observed, but there is another possibility. Consider an elliptical ring as shown in Fig. 1, composed of particles with perfectly aligned orbital axes. Particle 1 is on the outer edge of the ring at pericentre, particle 2 is on the inner edge, and orbit 2 has a slightly different eccentricity from orbit 1. The rest of the ring material is in orbits between 1 and 2. The orbits of particles 1 and 2 are described by the expressions for confocal ellipses:

$$\text{Orbit 1 } r_1 = r_{1p} (1 + e_1)/(1 + e_1 \cos \theta) \quad (1a)$$

$$\text{Orbit 2 } r_2 = r_{2p} (1 + e_2)/(1 + e_2 \cos \theta) \quad (1b)$$

where r_{1p} and r_{2p} are the pericentre distances, θ is the polar angle measured from pericentre, and e_1 , e_2 are the eccentricities. The ring width at pericentre is $\Delta r_p = r_{1p} - r_{2p}$.

The two points of the ϵ ring observed during the passage of SAO158687 differ in their distance to the centre of Uranus by

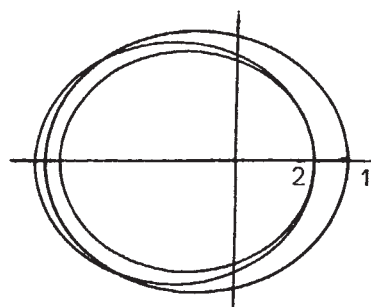


Fig. 1 Illustrating orbits at the edges of a ring. Two possible orbits for particle 2 are shown.

about 600 km. This sets a lower limit on the eccentricity of the ring of about $e = 0.008$. The two observed points, plus the assumption that the ϵ ring is never closer to Uranus than the next inner ring (the δ ring, at a radius of 48,400 km) imply an upper limit to the eccentricity of $e \approx 0.13$. This assumption should be valid if the rings are truly coplanar, but is suspect otherwise.

Equations (1a) and (1b) can be combined and expanded to first order in e to yield

$$\Delta r = r_1 - r_2 \approx \Delta r_p + \Delta r_p e_1 (1 - \cos \theta) - r_{1p} \Delta e (1 - \cos \theta) \quad (2)$$

where Δr is the radial spacing between the two orbits as a function of θ (as the orbits are not circles, this is not exactly the same as the width of the ring, but the difference is everywhere $< 10\%$ for $e \leq 0.4$), and $\Delta e = e_2 - e_1$.

If the term $\Delta r_p e_1 (1 - \cos \theta)$ in equation (2) is ignored (although it can make a significant contribution if e is as large as the 0.13 value mentioned above), we see that if $r_{1p} = 51,000$ km, Δe need be only as large as 0.0005 to give ring width variations of the order of 50 km. If $\Delta e \geq \Delta r_p / 2r_{1p}$, then $\Delta r = 0$ at at least one point around the ring. Presumably, a ring width of zero would be quickly broadened by interparticle collisions. How much broadening would occur and whether this applies to Uranus' rings is an open question.

According to equation (2), a broad variety of ring shapes are possible. The narrowest point can be at the pericentre ($\Delta e < 0$), or the apocentre ($0 < \Delta e \leq \Delta r_p / 2r_{1p}$), or anywhere in between ($\Delta e > \Delta r_p / 2r_{1p}$). In the last case, of course, there must be two narrow points located directly across the major axis from each other.

With only a single observation of the rings, the model cannot be appreciably constrained, but it does predict that ring width should vary smoothly around the ring in a way that can be fitted by equation 2. More observations, such as those suggested by Liller⁴, are eagerly awaited.

R. L. LUCKE

Code 685,
Goddard Space Flight Center,
Greenbelt, Maryland 20771

Received 21 November; accepted 14 December 1977.

1. Elliott, J. L., Dunham, E. & Mink, D. *Nature* **267**, 328-330 (1977).
2. Bhattacharyya, J. C. & Kuppuswamy, K. *Nature* **267**, 330 (1977).
3. Hubbard, W. B., Coyne, G. V., Gehrels, T., Smith, B. A. & Zellner, B. H. *Nature* **268**, 33-34 (1977).
4. Liller, Wm. *Astrophys. J.* **82**, 929 (1977).

λ -Transition between normal and superfluid ^4He in the high-speed rotating frame

THE rotating helium cryostat (1-m diameter, 3,000 r.p.m.) at Southampton provides a novel way of studying the behaviour of liquid helium. The centrifugal acceleration fields (up to 5,000g at the periphery) enable pressure differences up to 20 bar to be generated in radial ducts containing liquid helium between the axis and the periphery of the rotor. We have recently rotated the cryostat at temperatures down to 1.8 K, well below the λ -transition to superfluid helium. The λ -transition temperature is suppressed by increasing pressure so that the rotor becomes a unique tool for the study of the phase transition boundary between normal helium He(I) and superfluid helium He(II). Here we report observations of a step-like temperature difference of several tens of millikelvin across the boundary. The temperature jump can be explained in terms of the positive value of the volume coefficient of expansion for He(I) over a finite temperature interval above the λ -line, and the presence of a non-convecting intermediate state of finite thickness separating the two phases.

To consider the change in state during radial flow under rotation, the most convenient thermodynamic coordinates are the enthalpy (H) and the entropy (S). It can be shown that rotation increases the fluid enthalpy by an amount

$$\Delta H = \frac{1}{2} \omega^2 r^2 \quad \text{kJ kg}^{-1} \quad (1)$$

where ω is the rotation rate in rad s^{-1} and r the distance from the rotation axis¹. Previous work has demonstrated that in He(I), the change in state during radial flow is closely isentropic and the associated change in enthalpy is given by equation (1) to within the accuracy of the published thermodynamic data².

There is no published data in the form of an H - S diagram in the region of the λ -transition. However, using the available data, it was possible to generate an H - S diagram for helium in this region.

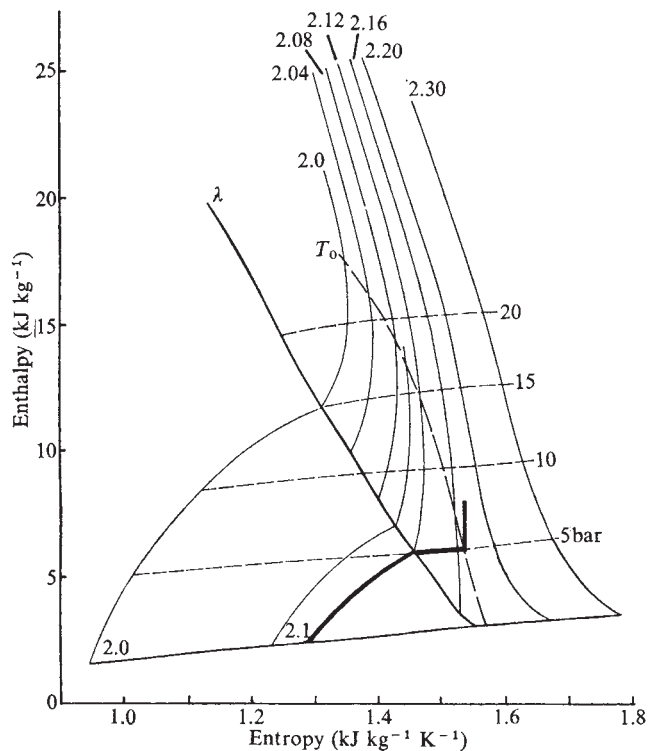


Fig. 1 H - S diagram for He^4 in the region of the λ -line derived from available data³⁻⁶. The dotted line labelled T_0 is the locus of the state described by $(\partial \rho / \partial T)_P = 0$. The heavy line describes the thermodynamic state along the radial duct from axis to periphery at 0.4 m and 2,500 r.p.m. at an axial temperature of 2.12 K when the He(I)/He(II) boundary is calculated to be at a radius of 0.34 m.

(He(II), ref. 3; λ -line, ref. 4; entropy and internal energy (U) of He(I), ref. 5; density-pressure relation, ref. 6. The enthalpy of He(I) was derived from $H = U + PV$.) Figure 1 gives the result of our calculations. The data were not evaluated critically so that this figure should be regarded as preliminary. The change in thermodynamic state with rotation is considerable. For example, with the present rotor 0.82 m diameter, operating at 2,500 r.p.m., the change in enthalpy is 5.5 kJ kg^{-1} between axis and periphery.

With axial temperatures below 2.10 K (that is well below the λ -line in the He(II) region), all the thermometers at different radial stations indicate the same temperature to within 10 mK. Hence an isothermal state exists along the radial duct from axis to periphery in spite of the considerable pressure gradient. This is to be expected because of the high effective thermal conductivity of He(II). This isothermal condition implies an entropy and superfluid density gradient along the radial ducts which is in accord with London's derivation of the mechano-caloric effect⁷.

For axial temperatures between 2.10 K and 2.17 K we have observed an easily controlled situation with He(I) in contact with

He(II) by a phase boundary between the two fluids. The coexistence of He(I) and He(II) in the Earth's gravitational field has been previously observed⁸. The radial location of the phase boundary can be predicted from the H - S diagram, with the He(II) in an isothermal state and the enthalpy increase due to rotation equal to $\frac{1}{2} \omega^2 r^2$. In practice, the boundary is easily moved in a radial direction by changing the rotation rate while the axial temperature is kept constant. The λ -transition is considered to be a second order phase transition according to Ehrenfest's classification⁹ since the free energy and entropy are continuous but the heat capacity is not. Usually in a second order phase transition, the two phases cannot coexist in equilibrium. In our case the coexistence between the phases is because of the imposed pressure gradient—a situation not explicitly considered by Ehrenfest.

The boundary region has an unexpected property, namely a significant positive step-like temperature difference between the He(I) and He(II) of several tens of millikelvins. This temperature jump is shown in Fig. 2 which shows the temperatures measured at radial stations as the He(I)/He(II) boundary was moved by changing the rotation rate. As this temperature jump was unexpected, experiments have been carried out to gain a physical understanding of this phenomenon and to eliminate the possibilities of experimental artefacts. The results are as follows: (1) at fixed axial temperature, ΔT is reproducible when the phase boundary is moved either inwards or outwards radially, that is no significant hysteresis effect was observed. (2) ΔT decreases as the axial temperature approaches the λ -temperature under the saturation vapour pressure (Fig. 3). (3) Within each of the two radial ducts, ΔT decreases with decreasing radius (Fig. 3), thus ΔT seems to be related to the local pressure gradient dP/dr as well as the intrinsic properties of helium. (4) In the presence of a large heat flux across the boundary applied by an electrical heater at the peripheral end of the radial duct, ΔT decreases non-linearly with increasing heat flux. For example at 40 mW cm^{-2} ΔT was 50% of its original value under the background heat flux of approximately 0.25 mW cm^{-2} . In other words, the temperature jump does not arise primarily from a Kapitza type thermal boundary resistance. (5) ΔT does not depend on self-heating (10^{-10} – 2.10^{-8} W) in the resistance thermometers or on differences in thermal resistance between the duct walls and He(I) or He(II).

In looking for an explanation of this temperature jump, it is important to realise that heat transfer in He(I) in the rotating

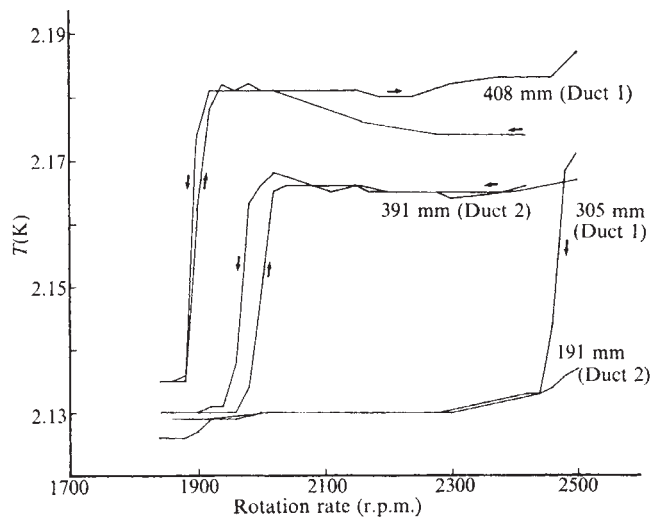


Fig. 2 Raw experimental data illustrating the temperature jump at the He(I)/He(II) phase boundary. The temperature at four radial stations, two in each duct, is plotted as a function of rotation rate with the axis temperature maintained constant at $2.125 \pm 0.005 \text{ K}$. The 391 mm (Duct 2) plot shows some hysteresis owing to rapid passage through the temperature jump. This hysteresis disappears when the rotation rate is changed more slowly as in the 408 mm (Duct 1) data. The experimental error on each temperature point is not more than $\pm 0.002 \text{ K}$.

frame is dominated by buoyancy driven convection processes in the high centrifugal acceleration field¹⁰.

We believe that the temperature jump can be explained in terms of the unusual temperature dependence of the density of He(I) just above the λ -transition. It is not solely the result of an intrinsic property of the He(I)/He(II) interface.

The radial dependence of the He(I) density within the duct can be written in the form

$$\frac{d\rho}{dr} = \left(\frac{\partial\rho}{\partial T}\right)_P \frac{dT}{dr} + \left(\frac{\partial\rho}{\partial P}\right)_T \frac{dP}{dr} \quad (2)$$

where it has been assumed that the temperature (volume expansivity) and pressure (isothermal compressibility) variables are separable. The temperature dependent term $(\partial\rho/\partial T)_P dT/dr$ is always negative but under adiabatic (isentropic conditions), it is much smaller than the pressure term, which is always positive, so that the density increases monotonically with radius.

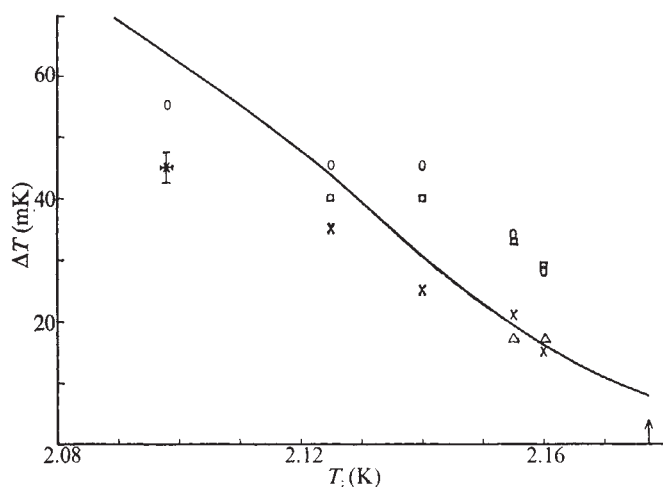


Fig. 3 Summary of all data, plotting the temperature jump ΔT as a function of T_λ . The solid line $T_0 - T_\lambda$ is derived from the published data. $\circ$, 408 mm, duct 1; $\square$, 305 mm, duct 1; $\times$, 391 mm, duct 2; $\triangle$, 191 mm, duct 2.

In the situation where heat is absorbed, buoyancy driven convection in the normal sense, that is upwards or radially inwards towards the axial heat sink, can only occur when $(\partial\rho/\partial T)_P$ is negative, regardless of its absolute magnitude in comparison with the pressure coefficient. In He(I), there is a region just above the λ -line where $(\partial\rho/\partial T)_P$ is positive and normal convection cannot take place.

If we restrict ourselves to the constant pressure situation (at constant radius) then we can examine the variation of $(\partial\rho/\partial T)_P$ in terms of the published data for $(\partial P/\partial T)_V$ and $(\partial P/\partial\rho)_T$. From Maxwell's equations

$$\left(\frac{\partial P}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial S}{\partial P}\right)_T \left(\frac{\partial P}{\partial V}\right)_T = -\left(\frac{\partial\rho}{\partial T}\right)_P \left(\frac{\partial P}{\partial\rho}\right)_T \quad (3)$$

Now $(\partial P/\partial\rho)_T$ is positive for both He(I) and He(II)¹¹ whilst $(\partial P/\partial T)_V$ is negative for $T_\lambda < T < T_0$ and positive for $T > T_0$. The interval $T_0 - T_\lambda$ increases with pressure and has a measured value of 70 mK at a pressure of 24 bar (ref. 12). From equation (3) it follows that $(\partial\rho/\partial T)_P$ is positive in the temperature interval between T_λ and T_0 and is zero at T_0 . In other words, under conditions of constant pressure, the temperature T_0 corresponds to the point of maximum density.

Alternatively we may derive $(\partial\rho/\partial T)_P$ again using Maxwell's equations thus:

$$\frac{1}{\rho} \left(\frac{\partial\rho}{\partial T}\right)_P = -\rho \left(\frac{\partial V}{\partial T}\right)_P = \rho \left(\frac{\partial S}{\partial P}\right)_T \quad (4)$$

From the H - S diagram, the behaviour of the isotherms can be used to indicate the variation of $1/\rho (\partial\rho/\partial T)_P$. Thus the locus of the point $(\partial S/\partial P)_T = 0$ describes the variation of T_0 with pressure, and is shown in Fig. 1. The coefficient $1/\rho (\partial\rho/\partial T)_P$ has a value of the order of 10^{-2} K^{-1} close to the λ -line. It should be noted that, $T_0 - T_\lambda$ is approximately 8 mK on the saturation vapour pressure curve, as reported elsewhere¹⁴.

In the presence of a heat flux through the wall of the radial duct containing a He(I)/He(II) interface at radius r_1 , the following situation will obtain. For $r < r_1$, the helium is superfluid with approximately constant temperature T_λ , determined by the vapour pressure at the He(II) surface near the axis, with heat being absorbed by vapourisation at the surface.

In the He(I) at $r > r_1$ when $(\partial\rho/\partial T)_P$ is positive, no convection towards the He(II) can take place. The effect of the heat flux through the wall at $r > r_1$ is to increase the temperature of the He(I) to the value T_0 at some radius $r_0 (> r_1)$ that is until $(\partial\rho/\partial T)_P$ is zero. Above this temperature when $(\partial\rho/\partial T)_P$ is negative, normal convection within the He(I) can take place. However, there is still a region adjacent to the interface which sustains the temperature difference $T_0 - T_\lambda$.

In this intermediate region, when $r_1 < r < r_0$, heat transfer cannot take place into the He(II) by convection in the normal sense. The temperature jump across this region is determined by the intrinsic properties of helium and to a first order of approximation is independent of the heat flux. Figure 3 shows that there is reasonable agreement between the predicted values of $T_0 - T_\lambda$ and the observed temperature jumps. In this curious situation the heat transfer mechanism across the intermediate region is not clear and may include contributions from, for example, conduction, oscillations or cellular fluid motion¹⁵, quantum mechanical tunnelling, and phonon mismatching.

If the heat transfer is by conduction alone, then the thickness of the intermediate layer is inversely proportional to the total heat flux through the layer. For example, in our experimental conditions, the thickness of this layer would be approximately 0.2 mm (compared with our experimental resolution of 4 mm determined by the size of thermometer) when the background heat flux was measured to be about 10^{-3} W . In these circumstances the isobaric approximation is reasonable. The time taken to establish the intermediate layer depends on the heat flux, but in our experiment it was generally shorter than the time taken to observe the temperature jump by, for example, changing the speed of rotation.

Note that the isobaric condition is not necessary for establishing the intermediate layer; only the condition that $(\partial\rho/\partial T)_P$ should be positive. It follows that, in the limit of zero heat flux, the intermediate state will extend through the He(I) out to the limiting radius r_0 where $(\partial\rho/\partial T)_P$ is zero. If r_0 is greater than the peripheral radius of the duct, all the He(I) will be in the intermediate state. In practice, the radial extent of the intermediate state will depend on the magnitude of the heat flux.

We have accounted therefore for the general features of the temperature profile in a radial duct containing the He(I)/He(II) transition. Further work is required to understand the detailed behaviour of helium in the presence of the intermediate layer. It must be emphasised, however, that neither our preliminary experimental observations, nor the accuracy of the available thermodynamic data on helium in the region of the λ -line are sufficient to enable a complete quantitative description to be made. A more detailed analysis of our findings will be presented elsewhere.

We thank our colleagues at Southampton for discussions on the thermodynamics of liquid helium, and A. Stevens for his

assistance with accurate temperature and speed control. This work was supported by an SRC contract under the UK Superconducting Generator Research programme.

R. G. SCURLOCK
D. B. UTTON

Department of Physics,
University of Southampton,
Southampton, UK

Received 21 November 1977; accepted 19 January 1978.

- Haseler, L. E., Scurlock, R. G. & Thornton, G. K. *Cryogenics* **16**, 331–336 (1976).
- Haseler, L. E., Scurlock, R. G. & Upton, D. B. *Cryogenics* **17**, 703–706 (1977).
- Maynard, J. *Phys. Rev.* **14B**, 3868–3891 (1976).
- Ahlers, G. *Phys. Rev.* **8B**, 530–568 (1973).
- Lounasmaa, O. V. *Cryogenics* **1**, 212–221 (1961).
- McCarty, R. D. *J. phys. Chem.* **2**, 923–1042 (1973).
- London, H. *Proc. R. Soc. A* **171**, 484 (1939).
- Ahlers, G. *Phys. Rev.* **171**, 275–282 (1968).
- Enrenfest, P. *Commun. Leiden Suppl.* no. 75b *Proc. R. Acad. Amsterdam* **36**, 153 (1933).
- Scurlock, R. G. & Thornton, G. K. *Int. J. Heat Mass Transfer* **20**, 31–40 (1977).
- Keesom, W. H. & Keesom, A. P. *Commun. Leiden Suppl.* no. 76B *Physica, Grav.* **1**, 128 (1933).
- Lounasmaa, O. V. & Kaunisto, L. *Ann. Acad. Sci. Fennicae AVI* no. 59 (1960).
- Ahlers, G. *Phys. Rev.* **8B**, 530–568 (1973).
- Wilks, J. *The Properties of Liquid & Solid Helium*, 296 (1967).
- Caspi, S., Lee, J. Y. & Frederking, T. H. K. *Cryogenic Engng Conf.* Boulder, Colorado (1977).

Dislocations at atomic scale

HIGH resolution electron microscopy has developed to such an extent that atomic scale details can now be observed: local information about atomic positions is therefore directly available^{1,2}. In contrast, diffraction experiments (X rays, neutrons, electrons) would give statistical data averaged over the whole sample, in addition to the difficulty of phase determination in complicated structures. The electron microscope is therefore a powerful and unique tool with which to study defects in the periodic structure of a crystal. We have applied this technique to dislocations and grain boundaries in germanium, and have determined their structure on a scale of 3–4 Å.

Owing to the strong electron interaction with matter, a one to one correspondence between the image and the projected crystal structure is restricted to axial illumination, symmetrical diffraction conditions, and perfect crystal conditions. But we have demonstrated that except in highly deformed regions (core region of 5 Å diameter) (ref. 3), this one to one correspondence is satisfied¹⁴. Tilted illumination introduces non-controlled artefacts⁴ and previous observations of dislocations^{5,6} must be considered with caution. Furthermore, the electron microscope aberrations limit the observable atomic plane distances to 3 Å at 100 keV. Thus, semiconductors with (111) atomic planes 3.25 Å apart are well suited for our study. The dislocation structure in semiconductors poses interesting questions: empty or full core structure as described by Hornstra⁷, dissociated or nondissociated cores⁸. We cannot definitely answer the first question, as no correct calculations are available for highly deformed regions. Therefore, we can only consider dissociated or nondissociated cores.

First, we investigated perfect germanium (110) thinned crystal. The relationship between atomic planes and the image was defined precisely both experimentally and theoretically⁹. For thicknesses less than 100 Å and certain defocusing distances (an easy parameter to control), the image looks like the projected structure of germanium at a resolution of 3 Å in a 100 keV electron microscope.

We then studied crystals containing defects in a suitable geometry. As the image is a two-dimensional projection along the whole crystal thickness, the simplest situation is obtained if a linear defect (dislocation) or a surface defect (grain boundary) is seen end-on, and if the displacement associated with the defect is constant along the beam direction. A pure tilt boundary in a bicrystal meets all these requirements if observed along the common axis. This method allows us to observe a great number of dislocations of the same geometry. Further-

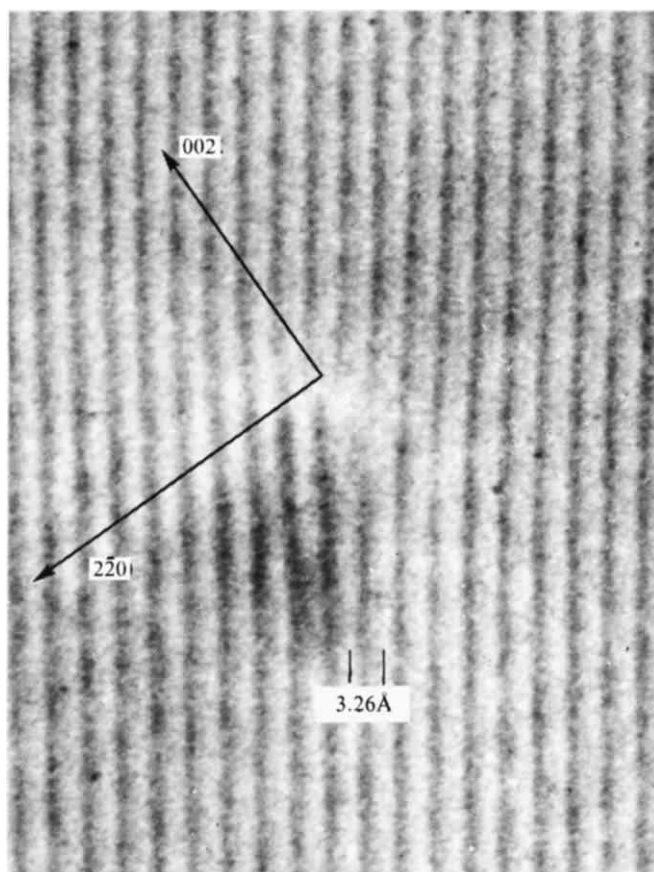


Fig. 1 A pure edge dislocation $b = 1/2[110]$ in a (110) germanium foil. (111) atomic planes seen with a 100 keV electron microscope (spherical aberration coefficient $C_s = 1.4$ mm). This dislocation belongs to a pure tilt low angle grain boundary (tilt angle $\theta = 0.5^\circ$).

more, by adjusting the tilt angle, θ , the distance, d , between dislocations is controlled: for a low θ value, dislocations are sufficiently far apart from each other to be considered independently (as far as the core structure is concerned) in contrast with medium or large θ values, where their interaction could be important. Tilting experiments performed on very thick regions have demonstrated that dislocations are aligned along the common axis at a precision of the order of $\pm 0.3^\circ$. Our results were obtained on [110] common axis bicrystals with θ values ranging from 0.5° up to 16° .

The $\theta = 0.5^\circ$ bicrystals¹⁰ contain pure edge dislocations or $b = 1/2[\bar{1}10]$ along [110], and 60° dislocations $b = 1/2[011]$ $1/2[\bar{1}01]$ associated by pairs so that the arrangement satisfies the Frank relation¹¹. The image of a pure edge dislocation (Fig. 1) exhibits the following features. First, as expected from the symmetry of the strain field associated with an edge dislocation, the image is not symmetric with respect to the (111) extra half plane. Computed atomic positions with an elastic model show a remarkable coincidence (± 0.3 Å) with experiments outside a 8 Å diameter central region where elasticity is no longer valuable. Second, the dislocation core is not extended. The radius given by a Peirls–Nabarro model¹² is too large to fit our experimental results. This can be understood physically because the observed edge dislocation is sessile and the covalently bonded germanium is a rigid structure. Thus we conclude that dislocations are not always dissociated in germanium.

Images of 60° dislocations have been similarly obtained: they exhibit a dissociation in two parts. The dissociation width is 30 ± 3 Å, from which we deduce a fault energy equal to 100 erg cm^{-2} .

The $\theta = 9.5^\circ$ bicrystals contained only pure edge dislocations separated by a mean distance $d = 24$ Å. As illustrated in Fig. 2, each dislocation is individually well defined and the two

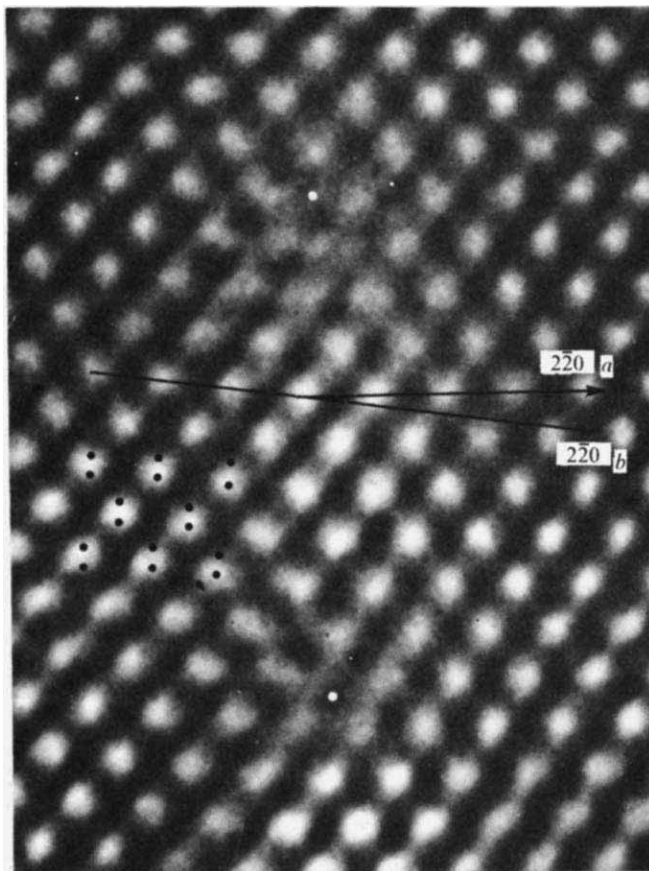


Fig. 2 Detail of a 9.5° pure tilt grain boundary in a (110) germanium foil: the projections of the atomic columns along [110] are visible (the perfect crystal cell is sketched). Dislocation cores are marked with white dots. 100 keV electron microscope. $C_s = 1.4$ mm. This picture has been optically treated to eliminate photographic and electronic noise without altering the structure details.

defects are separated by a good coincidence area. We observe the same behaviour at $\theta = 16^\circ$ ($d = 14$ Å). This confirms the Bollmann theory of primary dislocation network in a grain boundary¹³. However, we did not find any secondary dislocation: instead we observed that d fluctuates along the grain boundary between two values corresponding to the nearest coincidence orientations. This shows that in general a grain boundary is not strictly a periodic structure. Similar results have recently been observed on a high voltage electron microscope¹⁶.

Actually the resolution limit (3 Å) and the lack of good contrast calculations in highly distorted regions prevented our studying in more detail the core structure, or observing metals instead of semiconductors. Considerable difficulties remain in correctly interpreting images in the core region (5 Å diameter). However, several ways of going beyond this limit in defective crystals are being studied: the multislice approach to calculate dislocation core images¹⁵, the image treatment to correct distortions due to the aberrations, or the use of higher voltage electron microscopes.

A. BOURRET
J. DESSEAUX

Centre d'Etudes Nucléaires de Grenoble,
Section de Physique du Solide,
Département de Recherche Fondamentale,
85X, 83041 Grenoble Cédex, France

Received 16 August; accepted 20 December 1977.

1. Iijima, S. & Allpress, J. G. *Acta Crystallogr.* A30, 22–29 (1974).
2. Skarnulis, A. J., Iijima, S. & Cowley, J. M. *Acta Crystallogr.* A32, 799–805 (1976).
3. Bourret, A., Desseaux, J. & Renault, A. *Acta Crystallogr.* A31, 746–752 (1975).

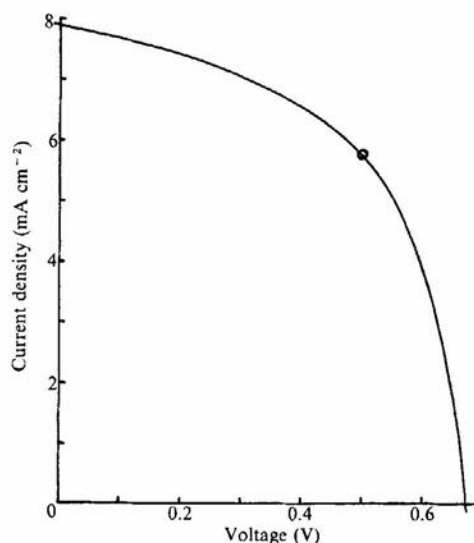
4. Cockayne, D. J. H., Parsons, J. R. & Hoelke, C. W. *Phil. Mag.* 24, 139–153 (1971).
5. Ray, I. L. F. & Cockayne, D. J. H. *Phil. Mag.* 22, 853–856 (1970).
6. Phillips, V. A. & Wagner, R. J. *J. appl. Phys.* 44, 4252–4254 (1973).
7. Hornstra, J. J. *Phys. Chem. Solid S*, 129–141 (1958).
8. Gomez, A., Cockayne, D. J. H., Hirsch, P. B. & Vitek, V. *Phil. Mag.* 31, 105–113 (1975).
9. Desseaux, J., Renault, A. & Bourret, A. *Phil. Mag.* 35, 357–372 (1977).
10. Bourret, A. & Desseaux, J. *J. Microsc. Spectros. Electron.* 2, 13–18 (1977).
11. Frank, F. C. *Rep. Symp. Plastic deformation of Crystalline Solids* 150 (Carnegie Institute of Technology, Pittsburgh, 1950).
12. Nabarro, F. R. N. *Proc. Phys. Soc. Lond.* 59, 256 (1947).
13. Bollmann, W. *Crystal defects and crystalline interfaces* (Springer, Berlin, 1970).
14. Bourret, A., Desseaux, J. & Renault, A. *J. Microsc. Spectros. Electron.* 2, 467–474 (1977).
15. Cowley, J. M. *Diffraction physics* (Elsevier, North Holland, 1975).
16. Krivanek, O. L., Isoda, S. & Kobayashi, K. *Phil. Mag.* 36, 931–940 (1977).

Amorphous silicon MIS solar cells

AMORPHOUS silicon grown by the rf glow discharge decomposition of silane in the process developed by Spear and Le Comber^{1,2} is one of the most promising materials found so far in the search for a cheap, large-area solar cell. Only two results, albeit encouraging, have been reported by RCA for p–n cells³ (AMI efficiency 2.4%) and anti-reflection coated platinum Schottky barrier cells⁴ (AMI efficiency 5.5%), but with area only $\leq 2 \times 10^{-2}$ cm². We report here the achievement of comparable efficiencies for a Schottky barrier cell, incorporating an insulating layer between the metal and semiconductor (an M–I–S cell) and with significantly greater area. The insulating layer is tailored to compensate for the low work function of the cheaper barrier metal (nickel) which we have used^{5–9}, leading to higher cell voltages otherwise obtainable only with higher work function metals. This is a significant step forward in the development of large area terrestrial solar cells.

We have observed efficiencies of 4.8% in 60 mW cm⁻² simulated sunlight, with a Ni–TiO_x contact on top of the amorphous silicon films, without an anti-reflection coating (Fig. 1). (That is, an estimated 6.3% with efficient anti-reflection treatment.) Diodes with 50% transparent contacts (areas up to 7×10^{-2} cm²) gave AMI short circuit densities of > 9.6 mA cm⁻², and open circuit voltages of > 600 mV. In 43 mW cm⁻² (actual sunlight) an open circuit voltage of 680 mV was given by the same diodes. This value of open circuit voltage compares extremely well with the typical value of 620 mV for a Pt barrier contact⁴. The fill factor was $\sim 51\%$, partially owing to the high series resistance of the amorphous silicon, although this is reduced from its very high dark value by incident AMI illumination.

Fig. 1 Current density–voltage relationship for M–I–S amorphous silicon solar cell, Ni–TiO_x barrier contact. Illumination, 60 mW cm⁻². Efficiency, 4.8%. $\circ$, $FF = 0.51$



The amorphous silicon, prepared at Dundee University, was undoped throughout most of its depth (1 μm) except for a narrow n^+ region adjacent to the stainless steel substrate to provide an ohmic contact. The top barrier contact was vacuum-deposited in an oil-free coating system immediately after the vacuum deposition of a thin film of TiO_x .

The open circuit voltage from these MIS diodes was a function of the insulator thickness, with an optimum between 15 and 25 Å. Neither the series resistance nor the reverse leakage current was increased by the addition of this layer unless it was thicker than 25 Å. With 20 Å of TiO_x beneath the Ni contact, the metal-semiconductor barrier was 1.1 eV high. Although nickel is not perhaps the ideal metal to use, it may be deposited by a number of methods, possibly including the rf discharge process.

The current density may be increased by adding an anti-reflection coating to increase the optical transmission through the barrier metal contact from ~50% to ~70%. The upper limit on this transmittance is set by reflections from the metal-semiconductor interface, rather than the air-metal interface. Further improvement in current is expected from an increase in the minority carrier lifetime in doped amorphous silicon. The short lifetime of holes in n-type amorphous silicon means that essentially only those created within the surface field region are collected by the metal contact. Those created by lower energy photons absorbed deeper within the cell are unable to diffuse to the field region before they recombine with electrons.

The devices remain stable when stored unprotected in air, and no degradation in performance has yet been observed. If the AMI short circuit current can be increased further to 18 mA cm^{-2} whilst still retaining an open circuit voltage of > 600 mV, and with an increased fill factor of 0.7, then even Ni contacts will give an efficiency in excess of 8%. Since the glow discharge process for depositing silicon films is readily amenable to scaling upwards, and since large area metallic and dielectric thin films prepared by vacuum deposition are already commercially available, the time is rapidly approaching when amorphous silicon MIS solar cells will be a commercial proposition.

We thank Professor W. E. Spear and Dr P. G. Le Comber for providing amorphous silicon films, and Professor S. D. Smith for suggesting this topic of amorphous silicon Schottky cells. This work was supported by the Wolfson Foundation and the SRC.

J. I. B. WILSON
J. MCGILL

Physics Department,
Heriot-Watt University,
Edinburgh,

S. KINMOND

Carnegie Laboratory of Physics,
Dundee University,
Dundee, UK

Received 10 November 1977; accepted 16 January 1978.

1. Spear, W. E. & Le Comber, P. G. *Solid State Commun.* **17**, 1193–1196 (1975).
2. Spear, W. E. & Le Comber, P. G. *Phil. Mag.* **33**, 935–949 (1976).
3. Carlson, D. E. & Wronski, C. R. *Appl. Phys. Lett.* **28**, 671–673 (1976).
4. Carlson, D. E. *IEEE Trans. Electron Devices* **ED-24**, 449–453 (1977).
5. Ponpon, J. P. & Siffert, P. *JAP.* **47**, 3248–3251 (1976).
6. Shewchun, J., Singh, R. & Green, M. A. *J. appl. Phys.* **48**, 765–770 (1977).
7. Landsberg, P. T. & Klimpe, C. *Proc. R. Soc. A354*, 101–118 (1977).
8. Viktorovitch, P. & Kamarinos, G. *J. appl. Phys.* **48**, 3060–3064 (1977).
9. Olsen, L. C. *Solid State Electron.* **20**, 741–751 (1977).

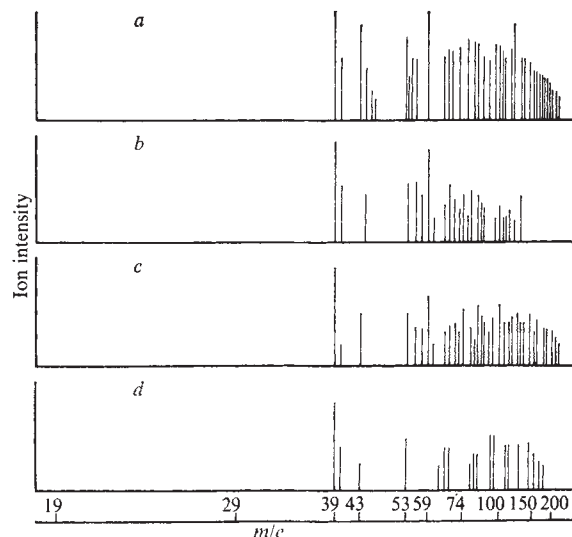


Fig. 1 Ion mass spectra of acetylene-oxygen flames. The ordinate represents ion intensities in an arbitrary logarithmic scale. The abscissa is linear with decreasing accelerating voltage of the mass spectrometer. *a*, Lean flame, equivalence ratio $r = 0.64$, height above burner top $h = 0.22$ cm, pressure $p = 21$ torr, initial velocity $v_0 = 62.5$ cm s^{-1} ; *b*, stoichiometric flame, $r = 1.03$, $h = 0.22$ cm, $p = 18$ torr, $v_0 = 69.9$ cm s^{-1} ; *c*, rich flame, $r = 2.05$, $h = 0.37$ cm, $p = 21$ torr, $v_0 = 58.6$ cm s^{-1} ; *d* as *c*, except $h = 11.23$ cm.

Ions have been observed in hydrocarbon pyrolysis in at least two laboratories^{2,3}, and Bowser and Weinberg³ attributed the source of ions to chemi-ionisation during pyrolysis. Chemiluminescence from excited CH and C_2 , normally observed in hydrocarbon flames, has also been observed in the infrared multiphoton induced pyrolysis of cyclopropane⁴. These species have long been known to be closely related to ionisation in flames^{5,6}.

Our experiments on oxy-acetylene were carried out in a 9.5 cm flat flame burner under sub-atmospheric pressures. Ions were monitored by particle beam sampling into a 360° magnetic sector mass spectrometer. Ion counts from a channeltron multiplier were fed into a multichannel analyser-accumulator. (The apparatus has been described elsewhere^{7,8}.) Possible errors due to sampling, estimated by the method of Hayhurst *et al.*^{9,10} could not affect our results significantly. The assignment of the $m/e = 39$ peak to C_3H_3^+ is supported by the intensities of the $m/e = 40$ peak (I_{40}) due to $^{13}\text{C}^{12}\text{C}_2\text{H}_3^+$. In lean and stoichiometric (but not in rich) flames, the intensities of the $m/e = 41$ peak (I_{41}) gradually increased and surpassed I_{40} in the burnt gas region when the sampling position above the burner top (h) is higher than about 0.9 cm, indicating some $^{41}\text{K}^+$. Under these circumstances, an upper limit correction was made on I_{39} , assuming that I_{41} was solely contributed to by $^{41}\text{K}^+$; the contribution of $^{39}\text{K}^+$ to I_{39} was then calculated from natural abundance data.

Figure 1 shows that initial ion mass spectra of lean, stoichiometric, and rich flames are very similar, and they in turn are similar to the spectrum of ions in the burnt gas high above the burner in a rich flame. The spectra show a prominent C_3H_3^+ peak and an abundance of heavy ions whose origin can only be attributed to pyrolysis reactions. Knewstubb and Sugden¹¹ observed a similar pattern and attributed it to fuel polymerisation in the reaction zone.

Figure 2 shows the profiles of relevant ions in lean, stoichiometric, and rich flames. The heavy ions with m/e between 75 and about 200 are grouped together in this figure. In lean and stoichiometric flames, these ions reach a maximum very early in the flame, then also disappear quickly. The next maximum belongs to C_3H_3^+ , which has long been suspected of being a primary ion¹. The maxima for CHO^+ ($m/e = 29$) and H_3O^+

Initial pyrolysis ions in hydrocarbon flames

CHEMI-IONISATION mechanisms in flames so far proposed have largely been based on the participation of oxygen¹. However, recent results from several laboratories and the results presented here strongly suggest that initial ions in hot hydrocarbon flames, whether lean, stoichiometric, or rich, may have a pyrolysis origin.

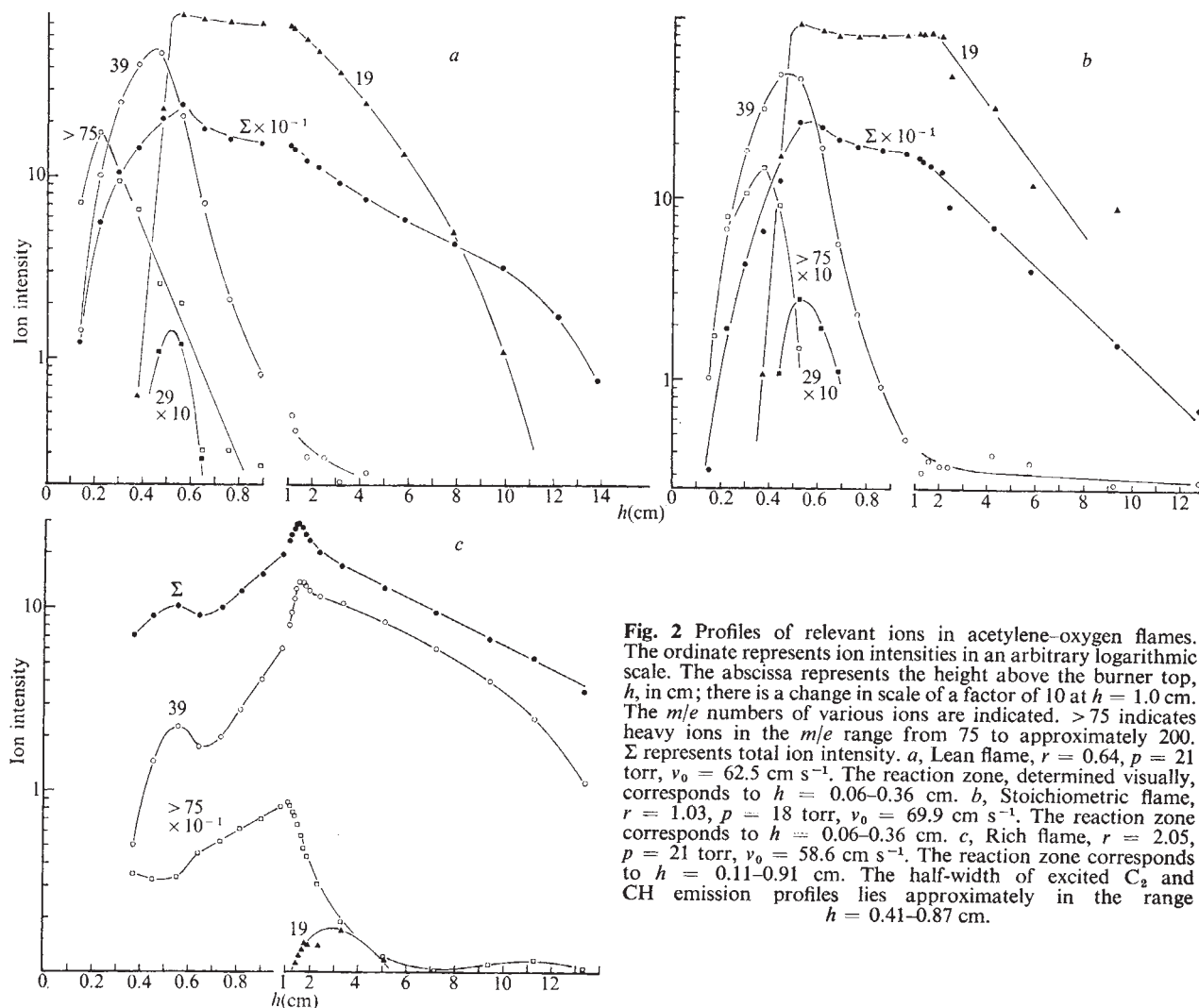


Fig. 2 Profiles of relevant ions in acetylene-oxygen flames. The ordinate represents ion intensities in an arbitrary logarithmic scale. The abscissa represents the height above the burner top, h , in cm; there is a change in scale of a factor of 10 at $h = 1.0$ cm. The m/e numbers of various ions are indicated. >75 indicates heavy ions in the m/e range from 75 to approximately 200. Σ represents total ion intensity. *a*, Lean flame, $r = 0.64$, $p = 21$ torr, $v_0 = 62.5$ cm s $^{-1}$. The reaction zone, determined visually, corresponds to $h = 0.06$ – 0.36 cm. *b*, Stoichiometric flame, $r = 1.03$, $p = 18$ torr, $v_0 = 69.9$ cm s $^{-1}$. The reaction zone corresponds to $h = 0.06$ – 0.36 cm. *c*, Rich flame, $r = 2.05$, $p = 21$ torr, $v_0 = 58.6$ cm s $^{-1}$. The reaction zone corresponds to $h = 0.11$ – 0.91 cm. The half-width of excited C_2 and CH emission profiles lies approximately in the range $h = 0.41$ – 0.87 cm.

($m/e = 19$) appear later, and much later than the maximum for heavy ions.

One significant aspect of Fig. 2 is that the $C_3H_3^+$ profile gradually resolves into two maxima as the equivalence ratio is increased; they are clearly resolved at an equivalence ratio of about 1.4 (unpublished data). These two maxima can be related to two zones in the flame, the pyrolysis zone and the oxidation zone, indicating that $C_3H_3^+$ ions are produced in both zones. The heavy ions in rich flames behave in a similar manner to $C_3H_3^+$.

When a fuel approaches a flame, it experiences a rapidly increasing temperature. At some distance before the position of maximum temperature, the temperature must be high enough for pyrolysis to take place. In fact, Tompkins and Long¹² observed polycyclic aromatic hydrocarbons in this region; the concentrations of these species reached a maximum very early in the flame, decreased slowly as oxidation took place, and then slowly rose to another maximum as soot began to form. We have observed what may probably be the corresponding ions. At the positions where the intensities of such ions reach a maximum in lean and stoichiometric flames, there must be very little CHO^+ ion (generally believed to be the major primary ion in flames¹) present, and the heavy ions are most unlikely products of proton or charge transfer from CHO^+ . The role of pyrolysis reactions in hydrocarbon oxidation has long received attention. For example, Bradley and Durden¹³ have shown that large quantities of oxygen have only a minimal effect on the basic pyrolytic decomposition and that hydrocarbon pyrolysis precedes the true oxidation. Our results are in agreement.

Our results and those of other workers suggest that the initial ions in flames may have a pyrolysis origin. What the primary ions are, by what mechanism they are produced, and how the disappearance of heavy ions is related to the appearance of smaller and oxygen-containing ions, would be major topics for further investigation. The suggestion of Bowser and Weinberg³ for the primary step



is plausible.

This work was done during R.S.T.'s sabbatical leave at the C.R.C.C.H.T. He thanks the C.N.R.S. for support and his French colleagues for their hospitality.

R. S. TSE*
P. MICHAUD
J. -L. DELFAU

Centre de Recherches sur la Chimie de la
Combustion et des Hautes Températures,
Centre National de la Recherche Scientifique,
45045 Orléans,
France

Received 1 December; accepted 22 December 1977.

*Present address: Department of Chemistry, University of Hong Kong, Pokfulam Road, Hong Kong.

1. Calcote, H. F. in *Ion Molecule Reactions* (ed. Franklin, J. L.) 2, 673–706 (Plenum, New York, 1972).
2. Abrahamson, J. & Kennedy, E. R. *12th Biennial Carbon Conf.* 169–70, Pittsburgh (1975).

3. Bowser, R. J. & Weinberg, F. J. *Combust. Flame* 27, 21–32 (1976).
4. Lesiecki, M. L. & Guillory, W. A. *J. chem. Phys.* 66, 4317–24 (1977).
5. Lawton, J. & Weinberg, F. J. *Electrical Aspects of Combustion* 255 (Clarendon, Oxford, 1969).
6. Kinbara, T. & Noda, K. *14th Symp. (int.) Combustion*, 321–28, Pittsburgh (1973).
7. Michaud, P., Delfau, J.-L., Barassin, A. & Thomas, R. *J. Phys. E* 8, 1030–32 (1975).
8. Delfau, J.-L., Michaud, P. & Barassin, A. *2nd Int. Symp. Plasma Chemistry*, Rome (1975).
9. Hayhurst, A. N., Kittelson, D. B. & Telford, N. R. *Combust. Flame* 28, 123–35 (1977).
10. Hayhurst, A. N. & Kittelson, D. B. *Combust. Flame* 28, 137–43 (1977).
11. Knewstubb, P. F. & Sugden, T. M. *7th Symp. (int.) Combustion* 247–53, Pittsburgh (1959).
12. Tompkins, E. E. & Long, R. *12th Symp. (int.) Combustion*, 625–34, Pittsburgh (1969).
13. Bradley, J. N. & Durden, D. A. *Combust. Flame* 19, 452–4 (1972).

Table 1 ^{10}Be analysis of Techno-1 manganese nodule

Sample depth (mm)	weight (g)	Chemical yield (%)	Net β count rate* (10^{-3} c.p.m.)	^{10}Be § (d.p.m. kg^{-1})
0–0.04	3.0	60	$14 \pm 8, 3 \pm 10, 10 \pm 6$	22 ± 10
0.04–0.2	5.4	66	$19 \pm 10, 19 \pm 7$	21 ± 7
0.2–0.7	32	{ 57 43†	$93 \pm 10, 91 \pm 9$ $60 \pm 10†, 64 \pm 12†$	20 ± 2 $18 \pm 3†$
0.7–2	70	20‡	$78 \pm 9, 75 \pm 10$	20 ± 5
2–4	95	19‡	$71 \pm 10, 76 \pm 16$	15 ± 4
4–7	186	{ 46 33†	$260 \pm 17, 261 \pm 15, 257 \pm 18$ $173 \pm 18†, 183 \pm 14†$	12.1 ± 1.2 $11.7 \pm 1.3†$
7–12	315	41	$145 \pm 17, 144 \pm 16$	4.5 ± 0.5
12–24	283	34	$28 \pm 8, 21 \pm 7, 26 \pm 11$	1.0 ± 0.2

* Background count rates were 0.283 ± 0.002 and 0.0070 ± 0.0003 c.p.m. for β - and α -channels respectively. Each ^{10}Be sample from the nodule was measured at least twice after an interval of ~ 4 weeks to check the constancy of its β activity. No α activity was found in any sample. The purity of samples was also checked by absorption measurements of β rays. The half-thicknesses of 22 ± 3 and 19 ± 4 mg cm^{-2} for 4–7 mm sample and 7–12 mm sample, respectively, are in good agreement with that of 21.8 ± 0.6 mg cm^{-2} for our ^{10}Be standard. Reference values were 21.2 ± 0.3 mg cm^{-2} (ref. 9) and 22.0 ± 0.7 mg cm^{-2} (ref. 16). In addition, 0.2–0.7 mm sample and 4–7 mm sample were subjected to further chemical purification and recounted. A good agreement was obtained between the results before and after the second purification. Chemical purity of BeO samples was checked by the scanning electron microprobe and any element of atomic number $z \geq 11$ was not detected. Blanks of the chemical reagents and the efficiency of the purification processes were checked by treating a near-shore sediment of 300 g, and no measurable activity was found for its BeO. These tests prove that the measured activities of the nodule samples are due to ^{10}Be .

† After second purification.

‡ Low chemical yields are due to the fact that a small amount of beryllium carrier (5 mg) was used for these samples, while that of 20 mg was used for the others.

§ Counting efficiencies were determined with ^{10}Be standards of known activity to be $27 \pm 2\%$ for 0.7–2 and 2–4 mm samples and $25 \pm 2\%$ for the others. The errors include all known sources of errors: counting statistics (1σ), counting efficiency, and chemical yield ($\pm 5\%$ except 0.7–2 and 2–4 mm samples: $\pm 20\%$).

Growth rate of manganese nodule measured with ^{10}Be and ^{26}Al

MANGANESE nodules are considered to have accumulated in general very slowly: surface layers of nodules contain a large excess of ^{230}Th , more than would be produced by uranium decay. The decrease of excess ^{230}Th with depth within a nodule is generally interpreted to be due to the radioactive decay, and nodule accumulation rates were then estimated from this to be a few mm Myr^{-1} (refs 1–4). How then do these nodules escape from burial by associated marine sediments which accumulate 3 orders of magnitude faster than the nodules? One possible explanation^{5,6} is that the radionuclides of interest were not incorporated in the nodule matrix during its growth but were adsorbed later. Subsequent inward diffusion of the radionuclides might result in an apparent radioactive decay. We have tested this hypothesis by measuring simultaneously two radionuclides of different half lives. If the gradients of the radionuclides are really due to the radioactive decay, then, assuming a constant growth rate for a nodule, we expect the profile of the shortlived nuclide to be steeper (because of its faster decay) than that of the longlived nuclide.

Bombardment of atmospheric constituents by galactic cosmic-rays produces the radionuclides ^{10}Be ($T_{1/2} = 1.5$ Myr) and ^{26}Al ($T_{1/2} = 0.716$ Myr) whose long half lives make them the most suitable for our study (refs 7, 8, respectively). These radionuclides are removed from the atmosphere chiefly by rain

and the portion which falls on the ocean eventually reaches the oceanic bottom^{9,10}.

The Techno-1 manganese nodule dredged at a depth of 4,020 m at the north of Tuamotu archipelago has a crust $\sim 1,000$ cm^2 in area. The oxide crust is ~ 2.5 cm thick. The underlying layer is a mixture of polymetallic oxides and covers a base of montmorillonite-rich assembly. For ^{10}Be measurements, half of the crust was scraped progressively in eight sections: 0–0.04, 0.04–0.2, 0.2–0.7, 0.7–2, 2–4, 4–7, 7–12, and 12–24 mm. This mode of sampling was chosen to find whether there was an excess of ^{10}Be near the surface or a simple exponential decrease. The excess would have been expected from the diffusion, but not from the radioactive decay hypothesis (experimental results later showed no such excess and confirmed the radioactive decay). ^{26}Al was measured in three sections: 0–2, 2–7, and 7–13 mm. The ^{26}Al content and the ^{10}Be content were uniform from the surface to the limit of the crust: 0.8% and 5 p.p.m., respectively.

The nodule material was leached in hot, 6 M HCl, and the insoluble phases were completely dissolved in HF. Beryllium carrier (5–20 mg) was added. Aluminium and beryllium were precipitated with NH_4OH . Iron, manganese and thorium were removed by the dissolution of $\text{Al}(\text{OH})_3$ and $\text{Be}(\text{OH})_2$ with NaOH. This purification process was repeated three more times. Beryllium was separated from aluminium by precipitation with NH_4OH in the presence of EDTA. Further purifications of beryllium were carried out by anion- and cation-exchange. The purified beryllium and aluminium were finally converted to BeO and Al_2O_3 .

^{10}Be is a pure β^- emitter with a maximum energy of 0.555 MeV (ref. 11). The purified BeO was counted with a gas flow type proportional counter. A pulse height analyser permits a distinction between β^- and α^- activities. The radiochemical

Fig. 1 Measured ^{10}Be - (a) and ^{26}Al - (b) concentrations as a function of depth in Techno-1 manganese nodule. The profile of shortlived ^{26}Al is steeper than that of longlived ^{10}Be , in agreement with the radioactive decay hypothesis. The absence of excess ^{10}Be activity at the surface confirms the radioactive decay hypothesis.

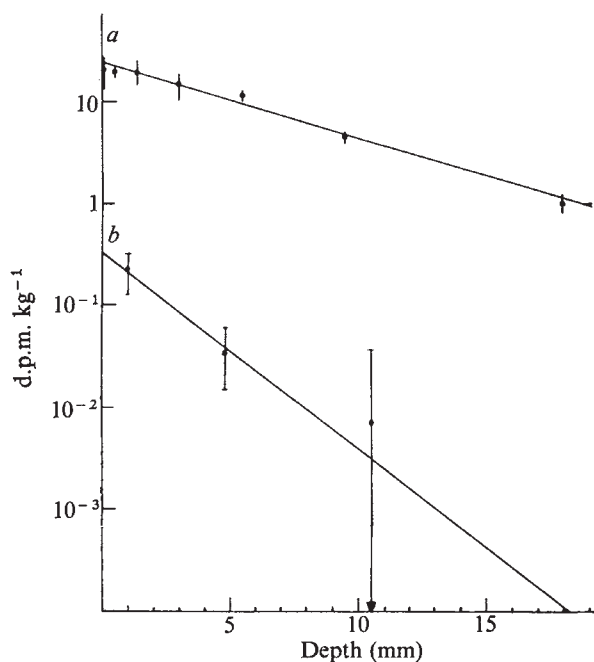


Table 2 Aluminium-26 analysis of Techno-1 manganese nodule

Sample Depth (mm)	Weight (g)	Net count rate* (10^{-4} c.p.m.)	$^{26}\text{Al}\dagger$ (d.p.m. kg^{-1})
0–2‡	148	15 ± 7	0.29 ± 0.14
0–2§	148	11.5 ± 8.5	0.18 ± 0.13
2–7§	621	9.6 ± 5.6	0.037 ± 0.022
7–13§	625	1.9 ± 7.3	0.007 ± 0.028

* The samples were counted for 20,000–60,000 min and the background for 90,000–140,000 min. The background count rates were 137 ± 5 times $\times 10^{-4}$ c.p.m. for 0–2 mm sample ‡ and $123.9 \pm 3.0 \times 10^{-4}$ c.p.m. for the others.

† The errors include all known sources of errors: counting statistics (1 σ), counting efficiency ($\pm 5\%$) and chemical yield ($\pm 10\%$). The counting efficiencies were 5.3% for 0–2 mm sample ‡ and 6.5% for the others. Chemical yield was 65%.

‡ Data reported in ref. 17.

§ Present work.

purity of the samples was carefully checked (see Table 1). ^{26}Al decays by β^+ emission (85%) followed by a 1.83-MeV γ -ray (99.7%) (ref. 11). The positron thermalises and produces two 0.511-MeV γ -rays, enabling the concentration of ^{26}Al to be measured by the coincidence detection of a 0.511-MeV γ -ray in one counter, and the sum peak of the other two γ -rays (total 2.34 MeV) in another. This coincidence peak is specific for the identification of ^{26}Al . The measurements of ^{26}Al were carried out with a low-level γ - γ coincidence spectrometer which consists of two semicylindrical NaI (Tl) scintillators (dimension 30 by 25 cm)¹².

Table 1 shows the results of the ^{10}Be measurements, and Table 2, those of the ^{26}Al . A ^{10}Be activity of 20 ± 3 disintegrations per minute (d.p.m.) per kg and a ^{26}Al activity of 0.23 ± 0.10 d.p.m. per kg were found in 0–2 mm section of this nodule, which corresponds to a $^{26}\text{Al}/^{10}\text{Be}$ ratio of 0.012 ± 0.005 . The value extrapolated at the surface is 0.014 ± 0.008 , in good agreement with those measured in oceanic sediment¹² and in Greenland ice¹³: 0.019 ± 0.012 and 0.017 ± 0.008 , respectively.

An exponential decrease with depth of ^{10}Be activity is evident for Techno-1 (Fig. 1). A growth rate of the nodule of 2.8 ± 0.6 mm Myr^{-1} was obtained by assuming that this decrease is due to the radioactive decay. A decrease of the ^{26}Al activities with depth is also apparent (Fig. 1). A nodule growth rate of 2.3 ± 1.0 mm Myr^{-1} was obtained from the ^{26}Al data, in good agreement with the ^{10}Be rate. These are comparable with rates measured by other methods on different nodules and also with the growth rates of three Pacific manganese nodules measured by Somayajulu¹⁴ and Krishnaswami *et al.*¹⁵ with the ^{10}Be analysis. The important point is that as ^{26}Al was measured from its specific peak by γ - γ coincidence spectrometer, the agreement of the ^{26}Al results with the ^{10}Be one in the present work reinforces the credibility of the ^{10}Be method. Much work must still be done before we thoroughly understand the origin of the manganese deposits. Nevertheless, the present studies with Myr cosmonuclides ^{10}Be and ^{26}Al fairly support the slow formation hypothesis.

We thank J. Labeyrie for his encouragement and C. Lalou and D'Ozouville (Centre National pour l'Exploitation des Océans) for the sample, also Nguyen Huu Van and C. Jehanno for atomic absorption analyses of Al and Be and scanning electron microprobe analyses. ^{230}Th and ^{231}Pa in the same nodule were measured by C. Lalou, E. Brichet and T. L. Ku, and their results will be published elsewhere.

F. GUICHARD
J.-L. REYSS
Y. YOKOYAMA

Centre des Faibles Radioactivités,
Laboratoire mixte CNRS-CEA,
91190-Gif-sur-Yvette, France

Received 30 November 1977; accepted 16 January 1978.

1. Goldberg, E. D. in *Oceanography* (ed. Sears, M.) 583–589 (A.A.A.S., Washington DC, 1961).
2. Bender, M. L., Ku, T. L. & Broecker, W. S. *Science* **151**, 325–328 (1966).

3. Barnes, S. S. & Dymond, J. R. *Nature* **213**, 1218–1219 (1967).
4. Ku, T. L. & Broecker, W. S. *Deep-Sea Res.* **16**, 625–637 (1969).
5. Arrhenius, G. *Trans. Am. Geophys. Un. U.S.A. natn. Rep.* 1963–1967, 604 (1967).
6. Lalou, C. & Brichet, E. C. *r. Acad. Sci. Paris D* **275**, 815–818 (1972).
7. Yiou, F. & Raisbeck, G. M. *Phys. Rev. Lett.* **29**, 372–375 (1972).
8. Samworth, E. A., Warburton, E. K. & Engelbertink, G. A. P. *Phys. Rev. C* **5**, 138–142 (1972).
9. Arnold, J. R. *Science* **124**, 584–585 (1956).
10. Goel, P. S. *et al. Deep Sea Res.* **4**, 202–210 (1957).
11. Lederer, C. M., Hollander, J. M. & Perlman, I. *Table of Isotopes* 6th edn (Wiley, New York, 1967).
12. Reyss, J.-L., Yokoyama, Y. & Tanaka, S. *Science* **193**, 1119–1121 (1976).
13. McCorkell, R., Fireman, E. L. & Langway, C. C., Jr *Science* **158**, 1690–1692 (1967).
14. Somayajulu, B. L. K. *Science* **156**, 1219–1220 (1967).
15. Krishnaswami, S., Somayajulu, B. L. K. & Moore, W. S. *Conf. Ferromanganese Deposits on Ocean Floor* (ed. Horn, D. R.) 117–121 (US NSF Washington DC, 1972).
16. Möler, P. J. *inorg. Nucl. Chem.* **32**, 2473–2481 (1970).
17. Reyss, J.-L. & Yokoyama, Y. *Nature* **262**, 203–204 (1976).

Excess ^3He and ^4He in Galapagos submarine hydrothermal waters

SINCE the discovery of excess ^3He in Pacific deep waters¹, it has been argued that the source of this anomaly is the mid-depth injection of primordial helium from seafloor spreading centres^{1,2}. This hypothesis is consistent with the spatial distribution of excess ^3He in the deep waters^{3–5}, its presence in the glassy (rapidly quenched) margins of extruded ocean basalts^{6,7}, and the detection of a large ^3He excess in a 'thermal plume' (thermal anomaly $\sim 0.1^\circ\text{C}$, sampled using a deep-tow sled) over the Galapagos Spreading Centre^{8–10}. In this report, we summarise 37 measurements of excess helium in hydrothermal waters sampled using the Alvin deep submersible at the Galapagos Spreading Centre during February and March of 1977.

The samples were taken from a series of four distinct vent fields using a continuous flow (pumped) system on the submersible which flushed an array of eight 8.8 l Plexiglass sampling bottles. The water was monitored using a CTD probe and numerous chemical measurements have been made on these samples. The temperature of waters sampled ranged from ambient (2.05°C) to 15.0°C , with a maximum uncertainty in the temperature of about 0.5°C .

The water samples were transferred on board ship to helium leak-tight copper tube samplers and returned for the shore-based helium analysis. The helium was extracted in a metal high vacuum system and stored in alumino-silicate glass sample tubes (Corning no. 1720). The isotopic ratio and amount of helium was measured by purifying and introducing a known fraction of the sample into a dual collection, statically operated helium isotope mass spectrometer^{3,4}. The helium contents were thus determined to $\pm 1\%$ and the isotopic ratio to $\pm 0.5\%$.

The helium found in the water is as much as 11 times enriched in ^4He , and 60 times enriched in ^3He relative to ambient seawater concentrations. All the samples cluster quite tightly about a straight line on a plot of ^3He against ^4He (Fig. 1), with an isotopic ratio ($^3\text{He}/^4\text{He}$) nearly eight times atmospheric. The black square in Fig. 1 represents ambient bottom water, that is water with helium contents typical of the area. These ambient waters also show an excess of ^3He and ^4He ($\sim 40\%$ and 10% of saturation, respectively) although much smaller than the effects observed here. The apparent scatter about the line (3.8%) is significant with respect to experimental error (0.5%); but the tightness of the correlation, despite the nearly fivefold range in absolute concentrations, suggests a common origin for the helium for all the vent fields. The hottest vents encountered, in the 'Garden of Eden' show the greatest scatter but with no significant bias from the general trend. This is significant in the light of the temperature data (discussed later). The isotopic ratio ($^3\text{He}/^4\text{He}$) of this excess helium ($1.08 \pm 0.02 \times 10^{-5}$) is in general agreement with the value $9.4 \pm 0.5 \times 10^{-6}$ observed by Lupton *et al.* using a deep-tow sled in a thermal plume ($\Delta T \sim 0.1^\circ\text{C}$) in this area (the uncertainty in their value we have estimated from their Fig. 1)^{8–10}.

The isotopic ratio is substantially lower than that found in tholeiitic glasses^{6,7} (1.3 – 1.7×10^{-5}). It is not possible to explain the higher isotopic ratios observed in the glasses by *in situ*

processes or partitioning between the water and rock. Diffusive loss from the glass after formation, partial equilibration with seawater, or *in situ* decay of U and Th all serve to lower the isotopic ratio of the glass helium. Further, this helium isotopic ratio is significantly lower than that observed in Red Sea brines¹¹ (1.2×10^{-5}), and distinctly lower than that of Kilauea gas, reported to be $2.09 \pm 0.11 \times 10^{-5}$ (ref. 4) and $1.94 \pm 0.04 \times 10^{-5}$ as measured in our laboratory. This suggests a significant degree of mantle heterogeneity (at least with regard to helium isotopes), which may be due either to a large scale variation in U and Th concentrations (and/or ^4He grow-in times) or to spatial inhomogeneity in the extent of mantle degassing. In either case, the measurement of helium isotopes should prove useful in studying mantle outgassing processes.

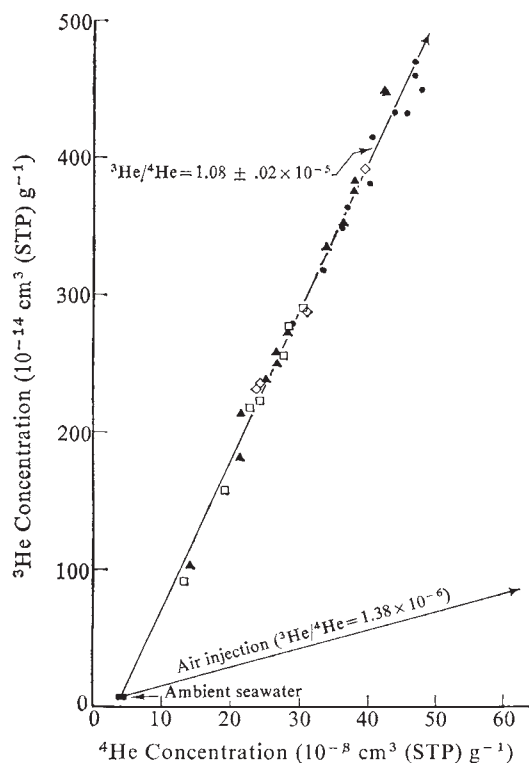


Fig. 1 The correlation between ^3He and ^4He for Galapagos hydrothermal waters. Ambient seawater (local bottom water, the black square) has ^3He and ^4He concentrations of 7.7×10^{-14} and $4.3 \times 10^{-8} \text{ cm}^3 (\text{STP}) \text{ g}^{-1}$ respectively. The air injection line represents the trend in concentrations that would result from the addition of atmospheric helium. The linear correlation coefficient for the linear regression line is 0.994. Vent areas: ●, Garden of Eden; ▲, Clambake; □, Oyster bed; ◇, Dandelion.

The ^3He concentrations measured correlate rather well with the temperature of the samples (Fig. 2). In fact, excluding the Garden of Eden samples, ^3He and temperature correlate linearly within errors ($\chi^2 \approx 1.9$), giving a slope of $7.6 \pm 0.5 \times 10^{-8} \text{ cal per atom}$. The Garden of Eden samples show consistent and significant bias from this correlation, appearing to be relatively rich in heat or poor in ^3He relative to the general trend. This is further mirrored by the isotopic ratio scatter (Fig. 1) and suggests that somewhat different processes occur in the vent field. Although not the only explanation, it is feasible that this effect may be due to an additional subsurface conductive transfer of heat to the circulating water, associated with little or no transfer of volatiles (that is, helium). It seems significant that this same vent shows anomalous behaviour (relative to the other vents) in a number of properties, including conductivity, alkalinity and pH. If the Garden of Eden data are included, the slope is changed to $9.5 \pm 0.7 \times 10^{-8} \text{ cal per atom}$, with a correspondingly increased χ^2 (6.2). In the light of other uncertainties in the calculations that follow, this difference

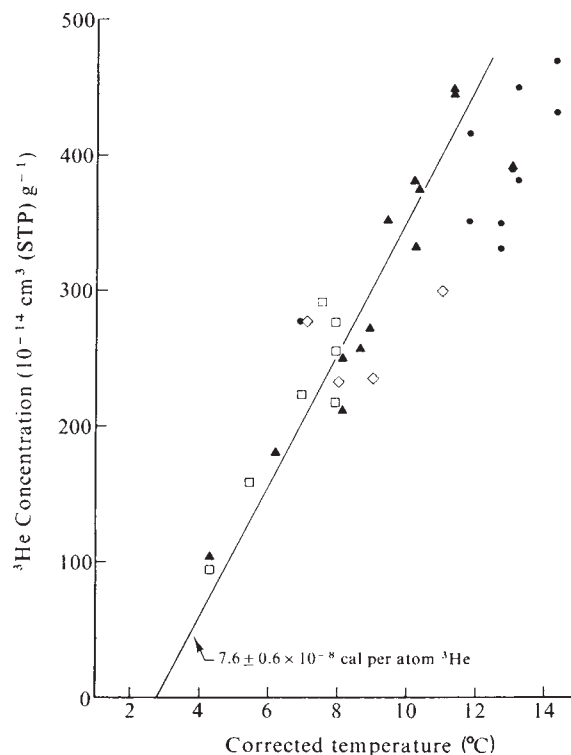


Fig. 2 The samples, exclusive of Garden of Eden, exhibit a tight correlation between excess ^3He and temperature ($\sigma(T) \approx 0.7^\circ\text{C}$), with a slope of $7.6 \times 10^{-8} \text{ cal per atom } ^3\text{He}$. Exclusive of the Garden of Eden samples (see text). Note the substantial and systematic offset of the Garden of Eden samples. Vent areas as in Fig. 1.

is only marginally significant. Regardless of the detailed mechanism, such a variation weakens the generalisations that we are about to make. However, until a more comprehensive survey is performed, our best approach will be to use the correlation obtained exclusive of the Garden samples, with the understanding that reinterpretation may be necessary later.

Because the ^3He anomaly has been detected throughout most of the world's ocean, and because (in the absence of tritium; as in the deep waters) helium is biologically and chemically 'conservative,' it should be possible to determine in a relatively unambiguous fashion its average oceanic flux. This estimate is further constrained by estimates of the ^3He loss rate in the upper atmosphere¹²⁻¹³ which must balance this flux (in addition to crustal and atmospheric production rates which are relatively small). The most recent estimate of this oceanic flux⁴ is $4 \pm 1 \text{ atoms cm}^{-2} \text{ s}^{-1}$ corresponding to $6.5 \times 10^{26} \text{ atoms yr}^{-1}$. Using the observed ^3He -heat correlation, we obtain a net 'convective hydrothermal heat flux' of $4.9 \pm 1.2 \times 10^{19} \text{ cal yr}^{-1}$. This value is in excellent agreement with the estimates of Wolery and Sleep¹⁴ ($4.0 \times 10^{19} \text{ cal yr}^{-1}$) and Williams and Von Herzen¹⁵ ($6.4 \times 10^{19} \text{ cal yr}^{-1}$) and represents the first independent confirmation of their calculations. Obviously, calculations of this sort may be extended to other (conservative) properties, thereby allowing an estimate of their ocean-mantle fluxes and enabling a more accurate balancing of their geochemical cycles.

It should be noted, however, that the ^3He -heat relationship gives a lower limit estimate to the hydrothermal flux as it is possible (in principle) to warm the circulating waters by heat conduction in the rock without exchange of helium, as perhaps suggested by the Garden of Eden results. Further, it must be established that the samples measured are indeed representative of all hydrothermal systems.

We have, therefore, detected a large component of excess helium in Galapagos hydrothermal waters with an isotopic ratio ($^3\text{He}/^4\text{He}$) substantially greater than atmospheric, but significantly smaller than primordial helium measured elsewhere,

further substantiating the hypothesis of mantle outgassing of helium, and indicating mantle heterogeneity. We have used an observed correlation between excess ^3He and temperature of the waters to compute a global hydrothermal convective heat flux of $4.9 \pm 1.2 \times 10^{19} \text{ cal yr}^{-1}$, which is in excellent agreement with theoretical estimates.

We thank M. Clauson, A. Bainbridge and the Alvin engineering group for the design of the sampling system and excellent technical support. This work was funded under NSF grant IDE 76-00389. Woods Hole Oceanographic Institution contribution No. 4035.

W. J. JENKINS

Department of Chemistry,
Woods Hole Oceanographic Institution,
Woods Hole, Massachusetts 02543

J. M. EDMOND

Department of Earth and Planetary Sciences,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

J. B. CORLISS

Department of Oceanography,
Oregon State University,
Corvallis, Oregon 97330

Received 22 September 1977; accepted 10 January 1978.

- Clarke, W. B., Beg, M. A. & Craig, H. *Earth planet. Sci. Lett.* **6**, 213 (1969).
- Craig, H. & Clarke, W. B. *Earth planet. Sci. Lett.* **9**, 45 (1970).
- Jenkins, W. J. thesis, McMaster Univ., Hamilton, (1974).
- Craig, H., Clarke, W. B. & Beg, M. A. *Earth planet. Sci. Lett.* **26**, 125, (1975).
- Jenkins, W. J. & Clarke, W. B. *Deep-Sea Res.* **23**, 481 (1975).
- Lupton, J. E. & Craig, H. *Earth planet. Sci. Lett.* **26**, 133 (1975).
- Craig, H. & Lupton, J. E. *Earth planet. Sci. Lett.* **31**, 369 (1976).
- Lupton, J. E., Weiss, R. F. & Craig, H. *Nature* **267**, 603 (1977).
- Weiss, R. F., Lonsdale, P., Lupton, J. E., Bainbridge, A. F. & Craig, H. *Nature*, **267**, 600 (1977).
- Lonsdale, P. *Deep-Sea Res.* **24**, 857 (1977).
- Lupton, J. E., Weiss, R. F. & Craig, H. *Nature* **266**, 244 (1977).
- Kockarts, G. & Nicolet, M. *Ann. Geophys.* **18**, 269 (1962).
- Johnson, H. E. & Axford, W. I. *J. geophys. Res.* **74**, 2433 (1969).
- Wolery, T. J. & Sleep, N. H. *J. Geol.* **84**, 249 (1976).
- Williams, D. L. & Von Herzen, R. P. *Geology* **32** (1974).

Evidence for an open seaway between northern and southern proto-Atlantic in Albian times

SINCE Wegener¹ presented his hypothesis on the origin of continents and oceans much effort has been centred on the dating of the first opening of a southern proto-Atlantic and on that of a first connection between the southern and northern Atlantic. Based on data from seafloor spreading, palaeomagnetism and the stratigraphy of the coastal basins flanking the Atlantic it was suggested that the opening of the southern Atlantic began in the Early Cretaceous²⁻³. According to the pattern of faunal distribution an open seaway with a free interchange of faunas should not have been established until the Lower Turonian⁴⁻⁵ though

some authors⁶⁻⁷ argue for an earlier opening in Albian times. We report here that new ammonite faunas from south-east Nigeria reveal a marine connection in Late Albian times.

The faunas were collected at three localities in the Calabar Province, south-east Nigeria, north of Calabar: road cuts along the main road north of Odukpani, the neostratotype⁸ of the Odukpani Formation; road cuts along the new road from Odukpani to Itu, and in the Mfamosing limestone quarry 25 km NE of Calabar. Here, along the south-east flank of the Oban Massif is the only place where marine Cenomanian rocks are exposed in Nigeria and in the Gulf of Guinea.

Sedimentation has been controlled by NW-SE trending faults in close relationship to the taphrogenesis of the SW-NE trending Benue trough, presumably a RRR junction⁹⁻¹¹ comparable to the Afar region of Ethiopia. Terrestrial sediments—the detritus from the Oban Massif—rest directly on the basement. These clastics of unknown age (Awi Formation) are succeeded by a sequence of sandy shales, shales with calcareous intercalations and limestones (Odukpani Formation) of Upper Albian to Lower Turonian age, deposited in marine shallow-water conditions. These marine shallow-water conditions persisted until the Late Campanian–Early Maestrichtian.

In the Mfamosing quarry Late Albian stromatolitic limestones with a thin condensed layer at the top are followed by pyritic shales of the Lower Cenomanian with an abrupt change in facies. This condensed horizon yielded a rich fauna of reworked ammonites of Uppermost Albian (Upper Vraconian) and Lower-Cenomanian age with abundant Late Vraconian Turritulidae (12%) and Early Lower Cenomanian Mantelliceratinae (61%). However, special weight has to be placed on the occurrence of the lyelliceratid genus *Salazicerias* (7%). This monotypic genus was known previously only by a small number of specimens, exclusively from the Lower Vraconian of Salazac, south-east France. Recently¹² it has been reported from the Lower Vraconian of the Bakony Mountains, Hungary, and from Tunisia. It has also been found in the Vraconian of the Tarfaya Basin, S-Morocco (J. Wiedmann, personal communication).

The geographic distribution of this genus—its limitation to the western Mediterranean, to Morocco and Nigeria—suggests a western interconnecting route. This is more probable than a long migration route along the East African coast, particularly considering the fact that rich Vraconian faunas are known from Madagascar, Mozambique and Zululand¹³⁻¹⁵, which contain the lyelliceratid genus *Stoliczkaia*, but not *Salazicerias*. There are no significant arguments for a direct 'trans-Saharan' passage at this time.

The western migration route is rendered plausible by the occurrence of marine Lower Cretaceous and Albian sediments off West Africa (Aaiun Basin, Senegal Basin)¹⁶⁻¹⁷, and of marine Albian in the Ivory Coast Basin¹⁸ and offshore Dahomey. The occurrence of *Dipoloceras*, *Neophlycticeras*, *Lyelliceras* and the endemic southern Atlantic genus *Elobiceras* in drilling cores from the Ivory Coast Basin¹⁸ even suggests an earlier, probable Middle Albian first connection. Furthermore the existence of a western seaway is evidenced by the associated ammonite fauna from Mfamosing quarry, and by Late Vraconian, Early to Middle

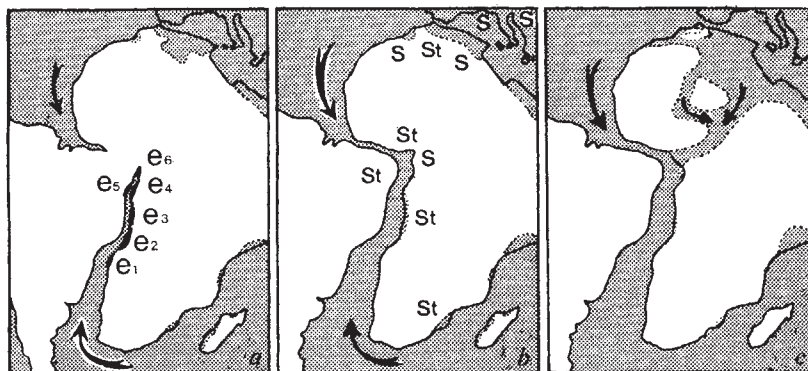


Fig. 1 Reconstruction of the relative positions of Africa and South America and of the palaeogeography of the South Atlantic during the Late Aptian (a), Late Albian to Early Cenomanian (b) and Late Lower Turonian (c). Occurrence of evaporites (e) in the basins of Mossamedes (e₁) (Lower Aptian), Cuanza (e₂) (Lower/Middle Aptian), Cabinda (e₃) (Middle Aptian), Gabon (e₄) (Late Aptian), Sergipe-Alagoas (e₅) (Upper Aptian–Lower Albian), and the Niger delta (e₆) (Lower Albian?). Distribution of representatives of the genus *Salazicerias* (S) (S-France, Hungary, Tunisia, S-Morocco and SE-Nigeria) and of *Stoliczkaia* of the *africana* (Pervinquière) group (St) (Tunisia, Algeria, Nigeria, Angola, Zululand and Brazil). The arrows indicate presumed routes of faunal migration.

Cenomanian faunas from south-east Nigeria and Angola⁶. There are close relationships between these and contemporaneous faunas from North Africa.

Palaeobiogeographic evidence is thus compatible with seafloor spreading and palaeomagnetic data. Initial rifting between Africa and America began during Triassic times in the north (Morocco–Senegal) and during Late Jurassic to Early Cretaceous in the south, followed by a sedimentation of a very similar pattern in the north and in the south. Sedimentation started with terrestrial sediments in graben structures, formed in the early stages of continental rupture. It was succeeded by Late Triassic to Early Jurassic evaporites in Morocco²⁰, the Aaiun Basin²¹, and by somewhat younger (Jurassic) evaporites in the Senegal Basin²² in the northern Atlantic rift system, and by Middle to Upper Aptian evaporites (Angola, Gabon, Niger delta, and Alagoas–Sergipe)^{11,23} in the younger southern rift system. The sources of the evaporites were oceanic water spilled over ridges into sub-sea level grabens. Regional subsidence of the continental margin caused marine incursions and the extension of epicontinental seas along the rift systems. Marine conditions have prevailed in the Morocco and Senegal basins since the Late Jurassic. The southwards transgressing sea probably reached the Ivory Coast Basin in Aptian times.

In the south a long, narrow sea extended as far as Gabon and Sergipe in Late Aptian times, evidenced by Aptian ammonites from the Walvis Ridge and from Angola⁷, Gabon²⁴ and Sergipe^{4,25}. By Middle Albian times, at the latest, the transgressing sea had reached south-east Nigeria. However, the rich Middle Albian *Oxytropidoceras* faunas from the southern proto-Atlantic show an endemic character. None of the Nigerian species seems to be conspecific with those of Brazil, and all faunas show little resemblance to contemporaneous *Oxytropidoceras* faunas from the Mediterranean or Texas. Associated ammonites known from the Tethys must be regarded as migrants along East Africa and the Cape seaway. The specific 'southern Atlantic' character with endemic mortoniceritids such as *Elobiceras*, *Angolaites* and *Neokentroceras* continued into the Early Upper Albian, though the occurrence of *Elobiceras* in the Ivory Coast Basin—the southernmost part of the northern rift system—supports a first connection between southern and northern proto-Atlantic in the Early Upper Albian. At the beginning this may have been more like a flooding of subsiding continental margin than actual rifting of the last barrier, though rifting has been reported in the Sierra Leone–Liberia area in Jurassic²⁶ and in Nigeria in Lower Cretaceous⁹ times. The Late Upper Albian and Early Lower Cenomanian fauna from the Mfamosing quarry and the Lower Cenomanian faunas from Odukpani present strong evidence of an open passage. One of the reasons why Tethyan ammonites occurred earlier in time in greater quantity and variety further to the south than southern Atlantic ammonites had migrated to the north might be the existence of unfavourable current systems in the early Atlantic.

REINHARD FÖRSTER

Bayer. Staatssammlung für Paläontologie,
Richard-Wagner-Strasse 10,
D-8000 München 2, FRG

Received 22 December 1977; accepted 19 January 1978.

- Wegener, A. *Die Entstehung der Kontinente und Ozeane* (Vieweg, Braunschweig, 1920).
- Larson, R. L. & Ladd, J. W. *Nature* **246**, 209–212 (1973).
- Herz, N. *Geol. Soc. Am. Bull.* **88**, 101–112 (1977).
- Reyment, R. A. & Tait, E. A. *Phil. Trans. R. Soc. B* **264**, 55–95 (1972).
- Offodile, M. E. *Publ. Paleont. Inst. Uppsala, spec. vol.* **4** (1976).
- Kennedy, W. J. & Cooper, M. J. *Geol. Soc. Lond. Trans.* **131**, 283–288 (1975).
- Wiedmann, J. & Neugebauer, J. *Init. Rep. DSDP* (in the press).
- Fayose, E. A. & Dessauvage, T. F. J. *Geol. Guide 7th African Micropaleont. Coll., Ile-Ife*, 35–52 (1976).
- Wright, J. B. in *Geology of Nigeria* (ed. Kogbe, X.) 309–317 (Eliz, Lagos, 1975).
- Burke K. & Whiteman, A. J. in *Implications of Continental Drift to the Earth Sciences* **2**, 735–755 (Tarling, D. H. & Runcorn, S. K.) (Academic London, 1973).
- Masclé, J. *Mem. Soc. Geol. France N.S.* **55** (128), (1977).
- Scholz, G. *Palaeontographica* (in the press).
- Collignon, M. *Atlas des fossiles caractéristiques de Madagascar*, **10** (Serv. Géol. Tananarive, 1963).
- Förster, R. *Geol. Jb.* **B 12**, (1975).
- Kennedy, W. J. & Klinger, H. C. *Bull. Brit. Mus. (nat. Hist.) Geol.* **25**, 265–315 (1975).
- Spengler, A. de, Castelain, J., Cauvin, J. & Leroy, M. in *Bassins Sédimentaires du Littoral Africain* (ed. Reyre, D.) 80–94 (Ass. Serv. Géol. Afr., Paris 1966).

- Schlee, J., Behrendt, J. C. & Robb, R. M. *Am. Assoc. pet. Geol. Bull.* **58**, 708–728 (1974).
- Spengler, A. de & Deltail, J. R. in *Bassins Sédimentaires du Littoral Africain* (ed. Reyre) 99–113 (Ass. Géol. Afr., Paris 1966).
- Förster, R. & Scholz, G. *N. Jb. Geol. Paläont.* (in the press).
- Faurer-Muret, A. & Choubert, G. in *Tectonics of Africa* (Unesco Earth Science Ser., 1971).
- Querol, R. in *Bassins Sédimentaires du Littoral Africain* (ed. Reyre) 27–39 (Ass. Serv. Géol. Afr., Paris 1966).
- Aymé, J. M. in *Salt Basins Around Africa* (ed. Ion, D.C.) 83–90 (London Inst. Petroleum, 1965).
- Burke, K. *Geology* **3**, 613–616 (1975).
- Klasz, I. de *Bur. Rech. Géol. Min.* **32**, 277–303 (1965).
- Beurien, K. Z. *dt. Geol. Ges.* **112**, 378–384 (1960).
- Jones, E. J. W. & Mgbatogu, C. C. S. *Nature* **267**, 688–690 (1977).

Record of last interglacial–glacial cycle from northeastern Queensland

DETAILED analyses of deep-sea sediment cores have produced spectacular advances in the knowledge of Quaternary stratigraphy and palaeoclimates. These have allowed the rigorous testing of hypotheses put forward to explain climatic fluctuations¹ and the construction of global atmospheric circulation models for dated periods in the past². Unfortunately, these studies do not provide evidence of the important climatic variable, precipitation, changes in which could be critical to the understanding of glacial advance and retreat and to atmospheric circulation patterns generally. Only sequences from terrestrial sediments can provide the necessary information on precipitation. Here again the majority of dry land sequences are of limited value (with certain exceptions^{3,4}) because they predominantly reflect temperature fluctuations or, where changes in effective precipitation are evident, sequences tend to be short or discontinuous. Here I outline and assess the significance of this sequence from tropical Australia which provides a long continuous precipitation record.

The record from Lynch's Crater covers about the past 120,000 years. The location of the site and results from the younger part of the sequence are detailed elsewhere^{5,6}. Selected results of pollen analysis of the full sequence are shown on Fig. 1. Four distinct periods can be recognised: zones E and A are dominated by pollen from a number of rain forest angiosperms; zone D contains high percentages of the rain forest gymnosperms *Araucaria* and *Podocarpus* and a significant sclerophyll component, while zone B is dominated by the sclerophyll taxa *Casuarina* and *Eucalyptus* comp; zone C is transitional between earlier rain forest and later sclerophyll assemblages. The total number of dry land taxa recorded—an indication of species diversity in the vegetation—is highest in the zones dominated by rain forest angiosperms and lowest within zone B, dominated by sclerophyll taxa.

Interpretation of the pollen diagram relies heavily on comparisons of the fossil spectra with spectra from surface samples collected within a variety of vegetation types in north-eastern Australia^{6,7}. Pollen spectra of zones E and A indicate the presence of rain forest existing under high rainfall conditions. Very low values of sclerophyll and rain forest gymnosperm taxa combined with high values for Cunoniaceae and *Elaeocarpus* and the significant presence of *Freycinetia* indicate rainfall peaks in subzones E3 and E1 and zone A. Within subzones E3 and E1, there is a general increase in *Elaeocarpus* relative to Cunoniaceae, a feature characteristic of the pattern of changes within Holocene pollen diagrams from this area⁸. Zone A and subzone E3 are very similar apart from the higher percentages of *Rapanea* in Zone A, but the lower values of *Freycinetia* and high percentages of *Cordyline* comp. in subzone E1 suggest that rainfall did not achieve similarly high levels. Effective rainfall was probably somewhat lower than today during the period represented by subzone E2.

A sharp reduction in rainfall is evident at the base of zone D, with marked increases in sclerophyll and rain forest gymnosperm pollen. Pollen of subzones D3 and D1 correspond with low araucarian vine forests which occur today in southern Queensland under a mean annual rainfall of between 900 and 1,200 mm (ref. 9). A higher rain forest angiosperm component and a peak in *Elaeocarpus* in subzone D2 indicate greater vegetation complexity under slightly higher rainfall.

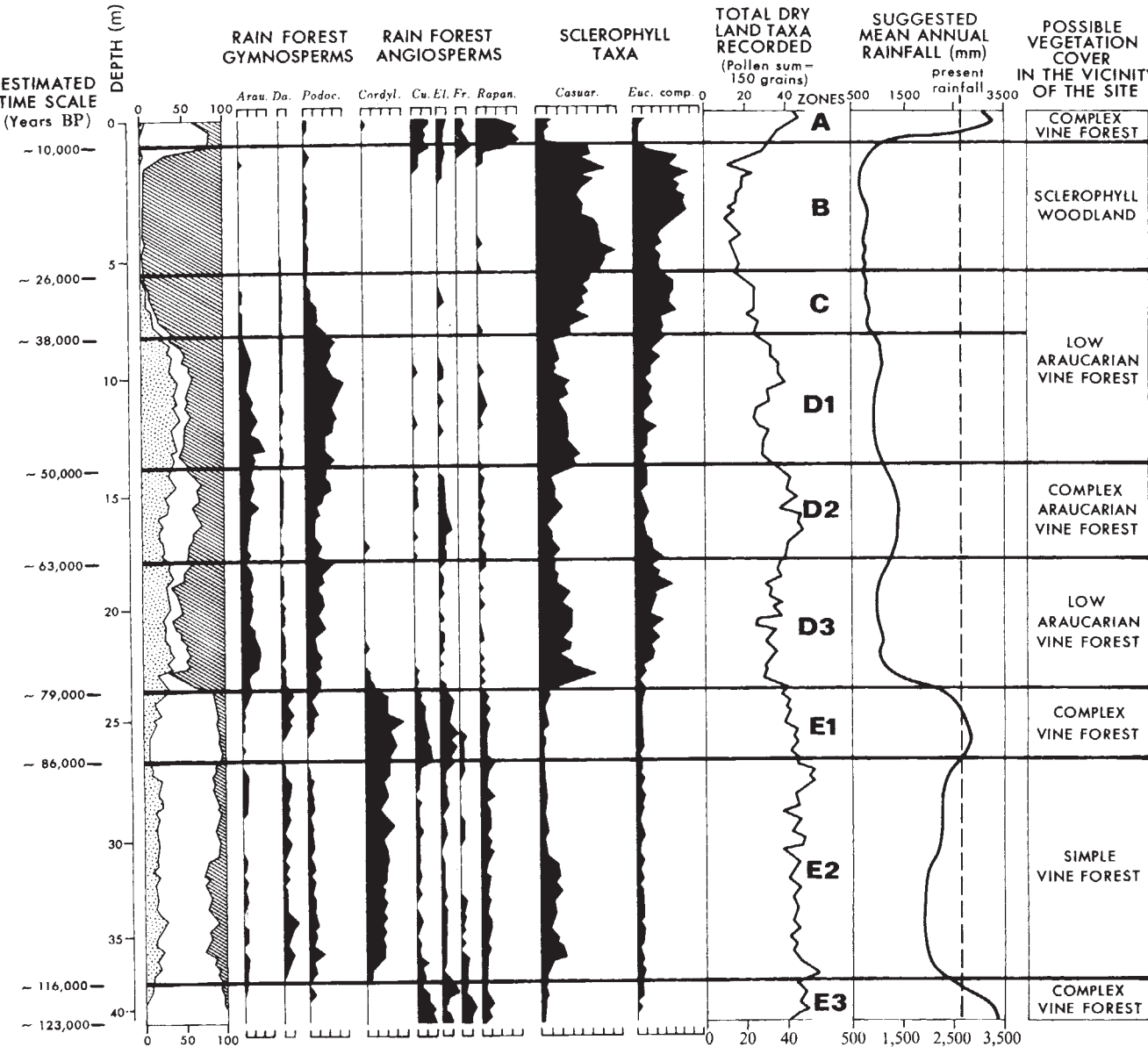
The replacement of araucarian forest by sclerophyll woodland within zone C is considered to be partly a result of a further decrease in effective rainfall and partly a result of burning by aboriginal man^{6,10}. Rain forest, which must have had a very restricted distribution during the zone B period, emerged from retreats in response to a sharp increase in rainfall at the beginning of the Holocene (zone A).

Significant evolutionary trends within the vegetation of the region have taken place during the period recorded in the diagram. *Dacrydium* and *Araucaria*, for example, which must have been important constituents of certain rain forest communities, did not recover after the dry sclerophyll period. Less dramatic changes have also occurred in the representation and composition of communities. A general and sustained increase in sclerophyll vegetation at the expense of drier rain forest types, probably a result of aboriginal man's activities, can explain some, but not all of the evolutionary changes shown.

Despite the importance of factors other than climate in the creation of vegetation changes, it has been possible to construct

a rainfall curve, the significance of which can only be adequately understood in relation to an absolute time scale. Such a scale has been determined for the past 40,000 years from radiocarbon dates, and estimated beyond this time by determining sediment accumulation rates of various parts of this core by comparison with known accumulation rates of similar sediments at this and other sites in the area, taking into account the degree of compaction of the sediment¹⁰. With the resultant time scale, a marked similarity between the rainfall curve and curves derived from analyses of deep sea cores is evident (Fig. 2). It has been previously suggested, from an examination of the upper part of this record, that rainfall changes within this area have been largely controlled by global changes in sea level and/or sea-surface temperatures^{13,14}. The extended record strengthens this suggestion. Correlations are particularly good between high rainfall peaks in subzones E3, E1, D2 and zone A and peaks in oxygen isotope substages and stages 5e, 5a, 3 and 1, respectively. There is no evidence of a rainfall peak corresponding with substage 5c, which is very prominent in curves denoting high

Fig. 1 Pollen diagram from Lynch's Crater. The column on the left shows the percentage of rain forest gymnosperms (stippled); rain forest angiosperms (white); sclerophyll taxa (hatched). The frequencies of pollen of all taxa are shown as percentages of the dry land plant pollen total; each division represents 10% of the pollen sum. Abbreviations for taxa; *Arau.*, *Araucaria*; *Da.*, *Dacrydium*; *Podoc.*, *Podocarpus*; *Cordyl.*, *Cordyline comp.*; *Cu.*, *Cunoniaceae*; *El.*, *Elaeocarpus*; *Fr.*, *Freyinetia*; *Rapan.*, *Rapanea*; *Casuar.*, *Casuarina*; *Euc. comp.*, *Eucalyptus comp.*



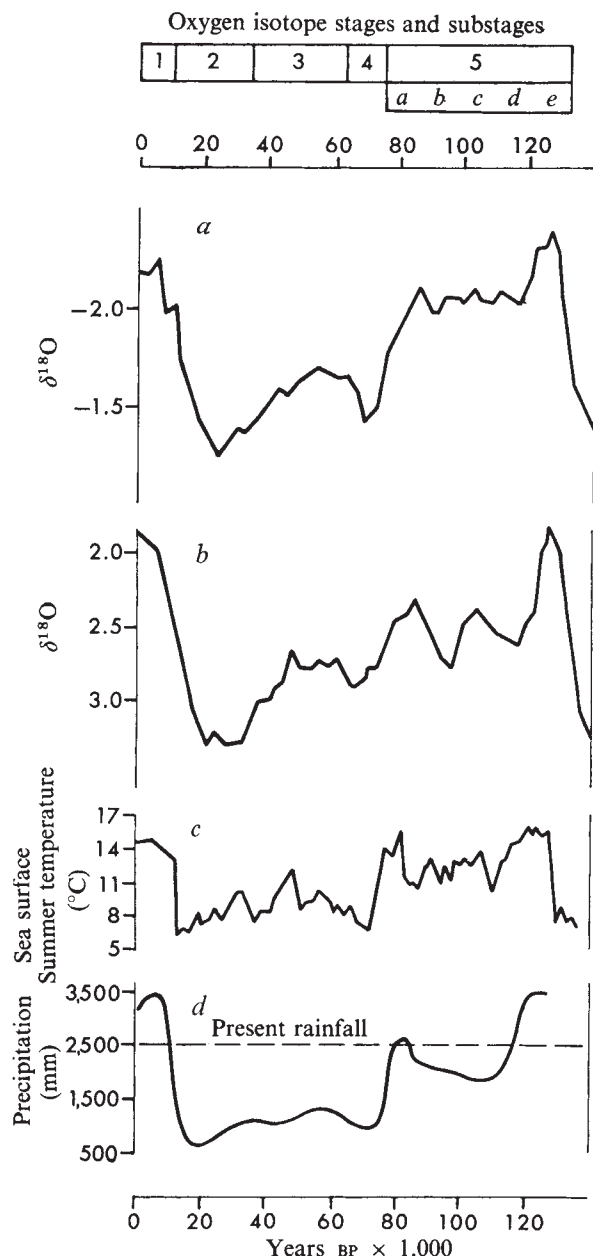


Fig. 2 The relationship between variations in rainfall around Lynch's Crater (d) and oxygen isotope and temperature curves from deep sea cores. a, Oxygen isotope curve from core V28-238 in the equatorial Pacific¹¹; b, oxygen isotope curve from core RC11-120 in the southern Indian Ocean¹; c, sea-surface summer temperature curve from core V23-82 in the North Atlantic¹².

sea-level phases from flights of raised coral reefs¹⁵ but not always identifiable or prominent in deep-sea core curves. The significance of the absence of this event in this detailed sequence is worthy of investigation.

I thank Dr G. S. Hope and C. Shaw who undertook the drilling of the older sediments; Professors D. Walker and B. G. Thom, Drs L. J. Webb and R. J. Wassen, and J. G. Tracey for help and advice, and A. and K. Anderson for permission to carry out field work on their property. The extended study was financed by an Australian Government Research Grant.

A. P. KERSHAW

Department of Geography,
Monash University,
Clayton, Victoria, Australia

Received 16 December 1977; accepted 12 January 1978.

1. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121-1132 (1976).

2. CLIMAP *Science* **191**, 1131-1137 (1976).
3. Wijnstra, T. A. *Acta Bot. Neerl.* **18**, 511-527 (1969).
4. Wijnstra, T. A. & Smit, A. *Acta Bot. Neerl.* **25**, 297-312 (1976).
5. Kershaw, A. P. *Nature* **251**, 222-223 (1974).
6. Kershaw, A. P. *New Phytol.* **77**, 469-498 (1976).
7. Kershaw, A. P. thesis, Australian National Univ. (1973).
8. Kershaw, A. P. *New Phytol.* **69**, 785-805 (1970); **70**, 669-681 (1971); **75**, 173-191 (1975).
9. Webb, L. J. & Tracey, J. G. *Austr. For.* **31**, 224-240 (1967).
10. Kershaw, A. P. *47th ANZAS Cong.* Hobart (1976).
11. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39-55 (1973).
12. Sancetta, C., Imbrie, J. & Kipp, N. G. *Quat. Res.* **3**, 110-116 (1973).
13. Nix, H. A. & Kalma, J. D. in *Bridge and Barrier: The natural and cultural history of Torres Strait* (ed. Walker, D.) 61-92 (Dept of Biogeography and Geomorphology. Publ. BG/3, A.N.U., Canberra, 1972).
14. Kershaw, A. P. & Nix, H. A. *Australian Conf. Climate and Climatic Change* Melbourne (1975).
15. Chappell, J. *Nature* **252**, 199-202 (1974).

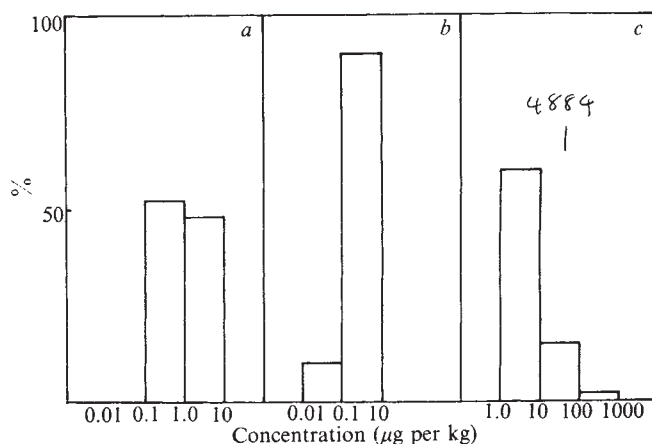
Estimate of the volatile nitrosamine content of UK food

MANY nitrosamines are carcinogenic¹ and can be formed by the interaction of nitrite with primary², secondary³ or tertiary amines⁴, or even quaternary ammonium salts⁵. Foods contain many amines, and in cured meats sodium nitrite provides the other precursor for the potential formation of nitrosamines. Some strains of bacteria reduce nitrate to nitrite, so that nitrate in conjunction with the appropriate amine could also lead to nitrosamine formation⁶. There have been several studies in which cured meats and a few other foods have been examined for nitrosamines, although most of these have been restricted to the detection of the volatile dialkyl and simple heterocyclic compounds⁷. Combined gas chromatography and mass spectrometry is the most reliable method, and has a typical detection limit of 1 µg per kg of food for volatile nitrosamines. More recently, the chemiluminescent detection of nitrosamines has enabled much lower concentrations to be measured⁸. We have carried out a survey of foods commonly encountered in the UK, using mass spectrometry⁹ and chemiluminescence, and we now summarise our data to assess the intake of volatile nitrosamines from these foods.

Foods were purchased in normal retail outlets in South-East England, and analysed within their shelf life, after cooking if appropriate. They were stored at 4°C or as normal for the product. We purchased bacon, canned luncheon and other cured meats, fresh meat and meat products, fish and fish products, cheese, yogurt, desserts, canned fruit and jams, frozen and fresh vegetables, soups, beverages and baby foods. Complete meals, prepared in domestic conditions, included various savoury flans, and grilled, fried and stewed meals. Meals were also prepared from commercial canned and other packaged complete meals.

We used our standard method of analysis¹⁰, in which food is slurried in water and the volatile constituents distilled. The distillate was extracted with dichloromethane and evaporated

Fig. 1 Occurrence of volatile nitrosamines in fried bacon (50 samples); a, N-nitrosodimethylamine; b, N-nitrosopiperidine; c, N-nitrosopyrrolidine.



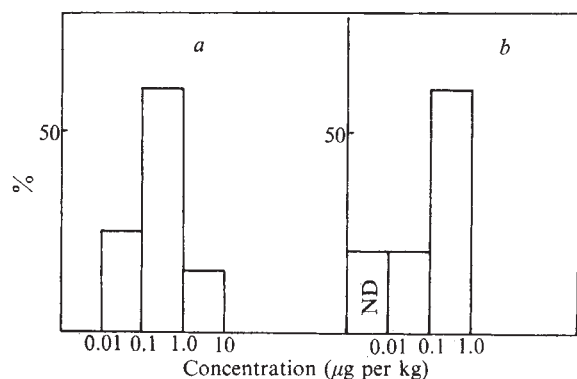


Fig. 2 Occurrence of volatile nitrosamines in other cured meats (75 samples); *a*, total volatile dialkyl nitrosamines; *b*, sum of *N*-nitrosopiperidine and *N*-nitrosopyrrolidine. ND, not detected, limit 0.01 µg per kg.

to a small volume to concentrate nitrosamines. The constituents of the extracts were separated by gas chromatography and detected by on-line high resolution mass spectrometry¹¹. Results were based on parent ion measurement after calibration against standard nitrosamine solutions. Each sample was examined for *N*-nitrosodimethylamine, *N*-nitrosodiethylamine, *N*-nitrosodi-*n*-propylamine, *N*-nitrosodi-*n*-butylamine, *N*-nitrosopiperidine and *N*-nitrosopyrrolidine. *N*-nitrosodi-*n*-propylamine was used as an internal standard for most samples so that recovery corrections could be made. The extracts were reexamined by chemiluminescence combined with gas chromatography to detect any of the above nitrosamines below the mass spectrometer detection limit, and any other volatile nitrosamines (including *N*-nitrosomorpholine) not listed above.

All samples of fried bacon contained *N*-nitrosopyrrolidine, in the range 1–20 µg per kg, but occasionally up to 200 µg per kg. *N*-nitrosopiperidine, previously not reported to occur consistently, was present in all samples between 0.08 µg per kg and 0.25 µg per kg. *N*-nitrosodimethylamine was also present in all samples, up to 5 µg per kg, and occasionally unidentified nitrosamines up to 0.5 µg per kg were detected. Figure 1 shows the frequency and range of occurrence of these nitrosamines in 50 samples. The concentrations are those in the cooked rashers as consumed, and do not include the greater proportions lost in the fat and by volatilisation during frying¹². The same nitrosamines occurred in other cured meats, but *N*-nitrosopyrrolidine did not exceed 1 µg per kg. Some cured meats simultaneously contained several nitrosamines at 0.1–1 µg per kg. Figure 2 gives the dialkyl and heterocyclic nitrosamine concentrations in 75 samples. In 36 samples of uncured meat (uncooked and cooked) and meat extract cubes, one sample contained 2 µg per kg of *N*-nitrosodimethylamine, and six samples contained dialkyl nitrosamines up to 0.2 µg per kg. Both uncooked and fried fish (Fig. 3*a*) contained *N*-nitrosodimethylamine, but out of 70 samples the only other nitrosamine found was *N*-nitrosopyrrolidine in one sample at 0.01 µg per kg. Salted, pickled, smoked and canned fish also contained on occasions *N*-nitrosodimethylamine (5 out of 24 samples). All common varieties of cheese available in the UK

were examined (Fig. 3*b*) and only *N*-nitrosodimethylamine was found, usually at low levels compared with the total nitrosamine content of cured meats. No nitrosamines were detected in 16 out of 76 samples. There was no more in toasted cheese than in the same cheese uncooked, nor in cheese with nitrate added, compared with varieties to which nitrate was not added during manufacture. Twenty-two other dairy products contained no heterocyclic nitrosamines, but three contained *N*-nitrosodimethylamine up to 0.1 µg per kg. Six canned desserts (baby foods) were all free of volatile nitrosamines above 0.01 µg per kg. *N*-nitrosodimethylamine was found in 5 out of 12 samples of canned fruit (< 0.1 µg per kg) and *N*-nitrosopyrrolidine in one sample (0.09 µg per kg). Sixteen samples of different varieties of vegetables contained no volatile nitrosamines. Three out of thirty soups contained *N*-nitrosodimethylamine (< 0.09 µg per kg) and two contained *N*-nitrosodi-*n*-butylamine (0.2, 0.5 µg per kg), one in the absence of any other nitrosamine. One sample contained *N*-nitrosopiperidine and *N*-nitrosopyrrolidine, both at 0.06 µg per kg.

Prepacked complete meals contained dialkyl nitrosamines on five occasions out of 13 samples (0.01–0.3 µg per kg). Ten canned baby food dinners were free of volatile nitrosamines, except for one containing bacon. The laboratory cooked meals consisted of 36 pastry based foods (including cured meat pies,

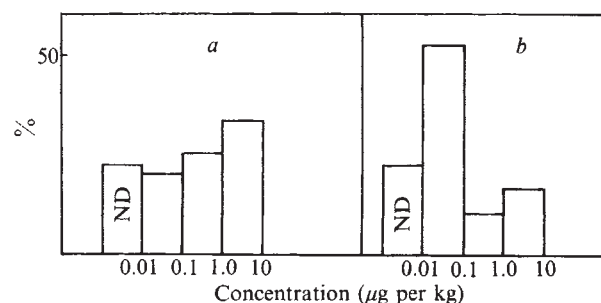


Fig. 3 Occurrence of volatile nitrosamines in fish and cheese; *a*, *N*-nitrosodimethylamine in fish (94 samples); *b*, *N*-nitrosodimethylamine in cheese (76 samples). ND, not detected, limit 0.01 µg per kg.

cheese and egg flans), 11 stews and casseroles, and 12 fried meals. *N*-nitrosodimethylamine was found in five of the pastry based foods (< 0.2 µg per kg), together with *N*-nitrosodiethylamine in one case, which was also present in two other samples (< 0.03 µg per kg). *N*-nitrosodimethylamine occurred on four occasions (< 0.5 µg per kg) and *N*-nitrosopyrrolidine once (0.08 µg per kg) in the other meals. *N*-nitrosoethylmethylamine occurred alone (0.2 µg per kg) and only in the four meals containing mushrooms.

Out of 493 samples analysed, cured meats represent the major source of volatile nitrosamines, followed by fish and cheese, and all other foods examined contained an average of not more than 0.06 µg per kg. The intake of these nitrosamines can be calculated from the average intake of the appropriate foods themselves¹³ (Table 1), assuming that all commonly consumed commodities likely to contain nitrosamines have been included in our survey. It is concluded that the likely intake of dialkyl nitrosamines per person from the normal UK diet is

Table 1 Estimated weekly intake of volatile nitrosamines from UK food

Food	Consumption (kg per week per person)	Dialkyl nitrosamines		Heterocyclic nitrosamines	
		average concentration (µg per kg)	intake (µg per week)	average concentration (µg per kg)	intake (µg per week)
Cured meats	0.34	1.0	0.34	8	2.7
Fish	0.14	0.2	0.03	—	—
Cheese	0.10	0.4	0.04	—	—
All other foods	9.62	0.06	0.58	—	—
Totals	10.20		0.99		2.7

approximately 1 µg per week and of volatile heterocyclic nitrosamines 3 µg per week.

We thank the National Cancer Institute for the loan of a Thermal Energy Analyser under contract NOICP 65795.

T. A. GOUGH
K. S. WEBB
R. F. COLEMAN

Laboratory of the Government Chemist,
Stamford Street,
London SE1, UK

Received 14 November 1977; accepted 24 January 1978.

1. Magee, P. N. & Barnes, J. M. *Br. J. Cancer* **10**, 114–122 (1956).
2. Warthesen, J. J., Scanlan, R. A., Bills, D. D. & Libbey, L. M. *J. Agric. Fd Chem.* **23**, 898–902 (1975).
3. Mirvish, S. S. *J. natn. Cancer Inst.* **44**, 633–639 (1970).
4. Smith, P. A. S. & Loeppky, R. N. *J. Am. chem. Soc.* **89**, 1147–1157 (1967).
5. Fiddler, W., Pensabene, J. W., Doerr, R. C. & Wasserman, A. E. *Nature* **236**, 307 (1972).
6. Coloe, P. J. & Hayward, N. J. *J. med. Microbiol.* **9**, 211–223 (1976).
7. Scanlan, R. A. *CRC Critical Reviews in Food Technology* **5**, 357–402 (1975).
8. Fine, D. H., Ruffeh, F., Lieb, D. & Rounbehler, D. P. *Analyt. Chem.* **47**, 1188–1190 (1975).
9. Gough, T. A., McPhail, M. F., Webb, K. S., Wood, B. J. & Coleman, R. F. *J. Sci. Fd Agric.* **28**, 345–351 (1977).
10. Goodhead, K. & Gough, T. A. *Fd Cosmet. Toxicol.* **13**, 307–312 (1975).
11. Gough, T. A. & Webb, K. S. *J. Chromat.* **79**, 57–63 (1973).
12. Gough, T. A., Goodhead, K. & Walters, C. L. *J. Sci. Fd Agric.* **27**, 181–185 (1976).
13. *Household Food Consumption and Expenditure for 1973* (HMSO, London, 1974)

Ultrasonic vocalisations facilitate sexual behaviour of female rats

ULTRASONIC vocalisations commonly occur during social interactions among rodents. During mating, adult male and female rats (*Rattus norvegicus*) emit brief 50–60 kHz ultrasonic calls¹; however, the function of these vocalisations is not known. In this study, we demonstrate that these mating calls have a precise function for communication. Specifically, 50-kHz vocalisations elicit sexual activity in female rats. Female rats exhibit a series of solicitation patterns during sexual behaviour, including orientation, darting and ear wiggling. These movements excite the male and enhance the likelihood of mounting, thereby facilitating copulation^{2–4}. (Beach has emphasised the importance of solicitation by the female during mating and has suggested the term 'proceptivity' to include the female's behaviour in the initiation and maintenance of copulation.) Although the presence of an intact male is usually a prerequisite for solicitation behaviour, the specific sensory cues which elicit it are not known. Isolated oestrous female rats exposed to ultrasonic vocalisations from a male exhibited a shorter latency to, and higher rate of darting when subsequently placed in a mating situation⁵. Although the auditory cues primed oestrous females to display increased solicitation behaviour when later confronted with an intact male, there was no indication that these ultrasonic vocalisations had a direct function for communication in the induction of these behaviour patterns. The objective of the present report was to determine if ultrasonic mating calls of rats have a direct function for communication during copulation.

Sixteen ovariectomised Long-Evans female rats served as subjects. Fifteen experienced males and 15 experienced female rats were used for transmitting vocal signals. Seven long-term castrated males were used as stimulus animals. The animals were maintained on a 12:12 h light-dark cycle (lights off at 0930). Ultrasonic vocalisations were electronically transmitted from a transmitting cage to a receiving cage. Both were glass-walled aquaria (52 × 26 × 29 cm high). The transmitting cage contained a central divider of wire mesh screen and was covered by a 10-cm thick sheet of polyfoam. The polyfoam was covered by corrugated cardboard within which a hole was made to hold a 0.64-cm Bruel and Kjaer microphone (model 4136). The microphone was placed directly over the male's section of the cage. Vocalisations were picked up using the microphone, and amplified, filtered and relayed by a transducer driver

amplifier to an electrostatic condenser microphone head connected for use as a speaker. The speaker was angled above the receiving cage. The speaker input was monitored on an oscilloscope and a Holgate ultrasonic receiver (bat detector) recorded vocalisations in the receiving cage. The transmitting cage and receiving cage were placed at diagonal corners of a 3 × 3 m room. Illumination was provided by red light above the receiving cage.

Female rats received subcutaneous injections of 50 µg oestradiol benzoate 54 h and 500 µg progesterone 6 h before the tests to induce behavioural oestrus. Testing was carried out during hours 4–8 of the 12-h dark period. Experimental females were placed in the empty receiving cage for a 2-min acclimation period. A castrated male was then introduced into the receiving cage and the stimulus treatments were begun simultaneously. Females were tested for 10 min in the receiving cage with one of the following treatments: (1) transmitted vocalisations from oestrous female and intact male previously given three intromissions (vocalisations, Voc); (2) transmitted sounds from empty cage (control, C); (3) transmitted sound from empty cage to receiving cage containing intact male urine (urine, U). (Urine from intact males was obtained on the test day by placing a male and oestrous female on opposite sides of a wire-mesh divider. The male typically urinated and his urine was collected on gauze squares. The urine-covered gauze was placed in a small wire box on the cage wall which enabled the female to smell the urine but prevented any gustatory contact.); (4) transmitted vocalisations from oestrous female and intact male to receiving cage containing intact male urine (urine+vocalisation, U+Voc). Each female received each treatment at 1-week intervals in a counterbalanced fashion. Darting and lordosis by the female in the receiving cage were manually scored on a pushbutton-activated Rustrak event recorder, as was the occurrence of transmitted vocalisations at 50 and 22 kHz. Darting and lordosis latencies and frequencies were calculated from the strip-chart record.

As shown in Table 1, females exposed to both treatments with vocalisation showed significantly more darting behaviour than those in the non-vocalisation treatments (two-way analysis of variance: $P < 0.01$). Similarly, both vocalisation treatments resulted in significantly shorter latencies to darting compared with control or urine alone treatments (two-way analysis of variance: $P < 0.01$). Table 1 also demonstrates that significantly higher lordosis frequently occurred during the two treatments involving vocalisations compared with non-vocalisation treatments (Friedman analysis of variance: $P < 0.001$). No differences were found between the Voc and Voc+U groups. It is important to note here that lordosis was usually exhibited in response to minimal contact stimulation by the castrated males. Castrated males were never observed to mount the females, and often during the vocalisation treatments a slight ano-genital sniff or nudge would be an adequate stimulus to elicit a lordosis posture. On occasion, lordoses were exhibited without any physical contact from the male.

It was generally observed that vocalisations occurred in an episodic fashion during the 10-min test period. The sexual responses of the female seemed to follow the fluctuations in vocalisation very closely. When transmitted vocalisations slowed or ceased, there was a decrease in the rate of solicitation; accordingly, when vocalisations were re-initiated, darting and lordosis were markedly increased in frequency. In some tests, males in the transmitting cage exhibited a transition from 50- to 22-kHz vocalisations. It was casually observed that at these times the females in the receiving cages immediately increased their rate of darting. A point-biserial correlation⁶ showed that there were significantly more darts in tests with (median, 78; $n = 8$) than in those without (median, 21.5; $n = 24$) 22-kHz vocalisations ($P < 0.001$).

The results of this experiment demonstrate that ultrasonic vocalisations are a sufficient cue to increase darting by a female rat in the presence of a castrated male rat. This comple-

Table 1 Proceptive behaviour of female rats in response to ultrasonic vocalisations

Behavioural measure	Control	Urine	Treatment group Vocalisation	Urine + vocalisation	P <
Darting (mean ± s.e.m.)	14.3 ± 3.0	16.6 ± 4.3	44.4 ± 8.8	41.0 ± 7.2	0.01*
Lordosis (median)	0.5	0.0	2.5	1.5	0.001†
Dart latency (s) (mean ± s.e.m.)	105 ± 48.8	154 ± 56.4	9.3 ± 2.4	15.3 ± 3.6	0.01*

*Two-way analysis of variance.

†Friedman two-way analysis of variance.

ments our earlier finding that vocalisations prime the female to dart when she is exposed to an intact male⁵. It seems that whether or not a male is present, vocalisations enhance proceptivity of females.

Previously, we have found that urine of males in the presence of vocalisations synergistically primed females to dart in subsequent mating with intact males. In the present study, the presence of male urine did not augment the effect of the transmitted vocalisations in the presence of a castrated male. This may reflect either the difference in the hormonal status of the stimulus male, or alternatively, whether or not the stimuli were presented in the presence of the stimulus male.

This experiment has shown that the threshold to display lordosis is reduced by ultrasonic vocalisations. To our knowledge, this is the first report of non-somatosensory cues augmenting lordosis frequency, an indication of receptivity. Detailed studies have shown how somatosensory stimulation of the female by the male initiates lordosis⁷⁻¹⁰. It would seem, therefore, that ultrasonic vocalisations are also significant sensory inputs which influence the initiation of lordosis.

Oestrous female hamsters exhibit lordosis in response to tape recorded vocalisations of the male¹¹. Similarly, partially deafened oestrous hamsters do not normally exhibit lordosis postures in the presence of a male with which they have no actual contact¹². In mice, vocalisations are correlated with the sexual readiness of the male and the presence of the odour of the female¹³; however, a causal relationship between male vocalisations and the response of the female has not yet been demonstrated in this species. On the basis of the above results and those presented here, it would seem that ultrasonic vocalisations are significant in the integration of mating activity in rodents in general.

In the present study, a changeover in the transmitted vocalisation from 50 to 22 kHz resulted in increased female solicitation behaviour. It has previously been shown that males which shift to 22-kHz vocalisations were most active in mating tests^{14,15}. It has also been shown that 22-kHz vocalisations sometimes occur during the copulatory sequence, particularly as ejaculation becomes imminent (ref. 16 and our unpublished observations). This suggests that 22-kHz vocalisations during the copulatory sequence may normally function to facilitate mating. In contrast, 22-kHz vocalisation following ejaculation are generally associated with sexual inactivity¹⁷. Twenty-two-kHz vocalisations also occur in a variety of other situations including solitary caging¹⁸, and subordination in agonistic encounters¹⁹. Twenty-two-kHz vocalisations in an agonistic context serve to inhibit attack²⁰, but in general the functional significance of these sounds is not known. It seems that the functions of 22-kHz vocalisations in social contexts is situation-dependent. Further study will be required to elucidate these functions.

TRACY K. MCINTOSH
RONALD J. BARFIELD
LYNETTE A. GEYER

Department of Biology,
Livingston College,
Rutgers University,
New Brunswick, New Jersey 08904

Received 14 October; accepted 28 December 1977.

1. Sales, G. D. *Anim. Behav.* 20, 88-100 (1972).
2. Hlinak, Z. & Madlafousek, J. *Physiologica Bohemoslovaca* 23, 146-147 (1974).
3. McClintock, M. K., thesis, Univ. Pennsylvania (1974).
4. Beach, F. A. *Horm. Behav.* 7, 105-138 (1976).
5. Geyer, L. A., McIntosh, T. K. & Barfield, R. J. *J. comp. physiol. Psychol.* (1978).
6. Bruning, J. L. & Kintz, B. L. *Computational Handbook of Statistics* (Scott Foresman, Glenview, Illinois, 1977).
7. Diakow, C. J. *comp. physiol. Psychol.* 88, 704-712 (1975).
8. Kow, L. M. & Pfaff, D. W. *Horm. Behav.* 6, 259-276 (1975).
9. Pfaff, D. & Lewis, C. *Horm. Behav.* 5, 317, 335 (1974).
10. Pfaff, D., Lewis, C., Diakow, C. & Keiner, M. in *Progress in Physiological Psychology* 1 (eds Stellar, E. & Sprague, J.) (Academic, New York, 1972).
11. Floody, O. R. & Pfaff, D. W. *J. comp. physiol. Psychol.* 91, 820-830 (1977).
12. Beach, F. A., Stern, B., Carmichael, M. & Ranson, E. *Behav. Biol.* 18, 473-487 (1976).
13. Whitney, G., Alpern, M., Dizinho, G. & Horowitz, G. *Anim. Learning Behav.* 2, 13-18 (1974).
14. Geyer, L. A. & Barfield, R. J. *J. comp. physiol. Psychol.* (1978).
15. Geyer, L. A., Barfield, R. J. & McIntosh, T. K. *J. comp. physiol. Psychol.* (1978).
16. Adler, N. J. *Eastern Conference on Reproduction Behavior*, Skidmore College, Saratoga Springs, New York, 13-16 June, 1976.
17. Barfield, R. J. & Geyer, L. A. *Science* 176, 1349-1350 (1972).
18. Francis, R. L. *Nature* 265, 236-238 (1977).
19. Sales, G. D. *J. Zool.* 168, 149-164 (1972).
20. Lore, R., Flannelly, K. & Farina, P. *Aggr. Behav.* 2, 175-181 (1976).

Ultradian rhythm of plasma noradrenaline in rhesus monkeys

KLEITMAN has postulated that a 'basic rest-activity cycle' exists which is an extension of the rhythmically recurring patterns of sleep cycles in man¹. Such ultradian rhythms, with a periodicity of approximately 90 min in non-human primates, have been associated with motor activity² and feeding behaviour³. Additionally, we have reported that plasma cortisol secretion in the rhesus monkey occurred in 90-min cycles which were synchronised for an entire group of monkeys over a 6-mth period, despite individual housing conditions of the animals⁴. Plasma noradrenaline (NA) is a reliable indicator of sympathetic nervous system activity⁵, in part because of its short half life of less than 2 min⁶. If Kleitman's hypothesis is correct, one might expect to see a basic 90-min rhythm for plasma NA, even though its short half life would favour the finding of a much faster rhythm. We have found a 90-min rhythm for plasma NA in the rhesus monkey, which was temporally synchronised for the experimental group as a whole. The presence of this rhythm suggests a tonic waxing and waning of sympathetic nervous system activity compatible with Kleitman's hypothesis.

Five adult male rhesus monkeys (*Macacca mulatta*) weighing 6.4-8.0 kg, were housed individually in primate restraining chairs within closed, sound-attenuating booths. Lights were turned on at 0700 and off at 1915. Behaviour was observed through a one-way mirror and feeding was carried out on a restricted 4-h schedule over 1300-1700. Water was available on an *ad libitum* basis. Blood was collected over a 4-h period at 15-min intervals through an intravenous or intra-arterial (monkey 3) catheter which extended outside the booth. Blood was collected (5 ml) in heparinised glass tubes and immediately spun at 2000g at 4 °C. Plasma was removed and frozen, and the red cells were resuspended in 5 ml 0.9% saline and re-infused immediately following the next blood drawing. Behavioural ratings were made just before each blood sampling. The state of arousal was noted according to whether the animal was asleep, tranquil, alert or excited. Blood samples were drawn over 0900-1300 in experiment 1 in all five animals and, 11-60 d after this, samples were drawn over 0900-1300 in monkeys 15 and 16, and over 1200-1600 in monkeys 3 and 12 in experiment 2. The latter two monkeys were fed on schedule at 1300 by opening the booth and manually placing the food within. Plasma samples

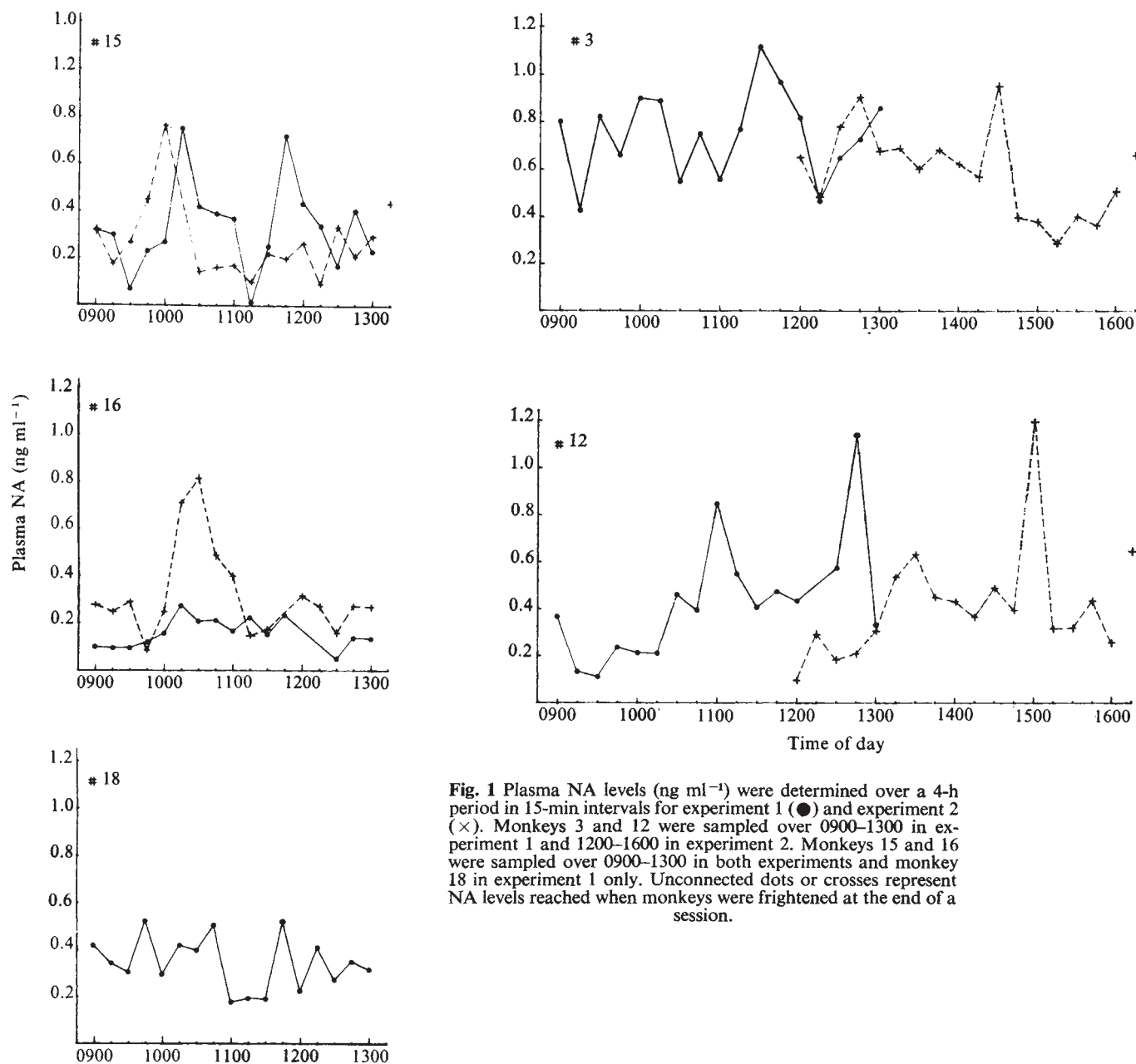


Fig. 1 Plasma NA levels (ng ml^{-1}) were determined over a 4-h period in 15-min intervals for experiment 1 (●) and experiment 2 (×). Monkeys 3 and 12 were sampled over 0900–1300 in experiment 1 and 1200–1600 in experiment 2. Monkeys 15 and 16 were sampled over 0900–1300 in both experiments and monkey 18 in experiment 1 only. Unconnected dots or crosses represent NA levels reached when monkeys were frightened at the end of a session.

were frozen at -70°C and NA levels were measured within 1 mth using the radioenzymatic method of Lake *et al.*⁶ which is sufficiently sensitive to measure $0.014 \text{ ng NA per ml plasma}$.

Plasma NA levels showed an ultradian rhythm in experiment 1 with peaks which were fairly well synchronised for the entire group of five monkeys, especially monkeys 3, 16 and 18 (Fig. 1). The data were normalised for the purpose of averaging and were found to show a bimodal pattern with large peaks 90 min apart at 1015 and 1145 for the group as a whole (Fig. 2), further suggesting synchronisation. The raw data from individual monkeys were each subjected to a smoothing procedure for best determination of peak values. The average interpeak interval for the five monkeys was $90 \pm 3 \text{ min}$ (s.e.m.) which was in good agreement with the period seen in the normalised, grouped data (Fig. 2). The normalised data were further dichotomised by the presence or absence of a peak plasma NA value greater than 1 s.d. above the mean during each of the consecutive 60-min epochs from 0900 to 1300; for example, monkey 15 in experiment 1 had one peak, respectively, in the 1000–1100 and 1100–1200 epochs. The distribution and synchronisation of these peaks was found to be nonrandom by an analysis of variance for repeated measures ($P < 0.05$) (ref. 7).

Repeated measurements of plasma NA at 0900–1300 in

monkeys 15 and 16 or at 1200–1600 in monkeys 3 and 12 in experiment 2, showed relative preservation of the periodicity in NA levels (Fig. 1). Average cycle length was $90 \pm 6 \text{ min}$ with variations of 15 min from this interval only in monkeys 3 and 12. Minor differences in phase synchronisation (15 min) were present in monkeys 3, 15 and 16 between experiments 1 and 2, despite the relative preservation of cycle length. Monkey 12 was, however, 45 min out of phase (Fig. 1).

Peak levels were often two- to threefold higher than the 'baseline' levels and occasionally rose to eight times these levels. There was no apparent correlation of the rhythmic alterations in NA levels with levels of arousal or feeding behaviour. Each animal was frightened by loud noise and threatening gestures just after the last blood sampling, and an additional sample was obtained immediately following this procedure. This was always associated with an increase in plasma NA above the final sample level, although the magnitude of this increase was not always as great as NA levels reached in the intrinsic, rhythmic peaks of NA during the preceding 4 h (Fig. 1).

We conclude that plasma NA levels vary in approximately 90-min cycles which are unassociated with levels of arousal or feeding behaviour. This rhythm was synchronised across the entire group of animals despite their isolation from each other.

B14 sublines⁶, the identified recombinant chicken could not have resulted from a 'mix-up' in our breeding colony. Further mating between M1^a/M1^b, G1^a/G1^a birds resulted in Mendelian-type segregation of M1 alleles in the F4 progeny and the chickens homozygous for the recombinant M1^b-G1^a chromosome have been designated as the new B14D strain.

Previous searches for intra-C_H gene recombination in mice which yielded negative results concerned only γ -chain allotype markers^{2,3}. Our finding of C γ -C μ crossing-over implies that the distance between C μ and the C γ gene in chickens is probably greater than the distance between C_H genes for different C γ isotypes in mammals. Since subclasses of chicken IgG so far have not been identified (G1 allotype is a marker for the bulk of IgG molecules⁴) it seems likely that the tight clusters of murine C γ genes diversified only at a later, mammalian stage of phylogenetic development.

Further studies comparing V_H gene products between the parental B14A, B14C strains and of the recombinant B14D strain may provide useful information on the genetic mapping of C_H and V_H genes. It has been reported recently⁷ that two mouse recombinants between gene markers for V_H and C γ , C α did not involve C μ markers. This suggests that C μ genes are more closely linked to V_H genes than they are to C_H genes for C γ or C α isotypes. Thus, analysis of the effects of the M1^b-G1^a cross-over event on the expression of V_H genes should provide pertinent information on the relative distance between V_H and C μ or C γ genes in chickens.

I thank Dr R. Pink for advice in the design of the genetic protocol, Mr David Ward for looking after the breeding flocks and Mr C. Makings and Ms L. Strudwick for technical assistance.

J. IVANYI

Department of Experimental Immunobiology,
Wellcome Research Laboratories,
Beckenham, Kent, UK

Received 4 November 1977; accepted 17 January 1978.

1. Weigert, M. & Potter, M. *Immunogenetics* 4, 401-435 (1977).
2. Herzenberg, L. A., McDevitt, H. O. & Herzenberg, L. A. *Rev. Genet.* 2, 209-244 (1968).
3. Mage, R., Lieberman, R., Potter, M. & Terry, W. D. in *The Antigens* (ed. Sela, M.) 1, 300-377 (Academic, New York and London, 1973).
4. Ivanyi, J. *J. Immunogenetics* 2, 69-78 (1975).
5. Pink, J. R. L. & Ivanyi, J. *Eur. J. Immun.* 5, 506-507 (1975).
6. Ivanyi, J. & Lydyard, P. *Immunogenetics* 2, 285-289 (1975).
7. Mage, R. *Prog. in Immunol., III Int. Congr. Immun.* (in the press).

Resistance of morphine-seeking behaviour in rats to pharmacological and behavioural 'treatments'

OPIATE addicts may be treated by the controlled prescription of an opioid drug such as methadone^{1,2}; such 'replacement' therapy aims to meet the addict's need for his drug and so to reduce self-administration. Treatment with antagonists is an alternative; this treatment is designed to diminish the efficacy of any self-administered opiate by 'blocking' the receptors with substances such as naloxone or naltrexone³. A third method of treatment involves behavioural aversion procedures, which aim to suppress the actions culminating in drug-taking by associating them with unpleasant stimuli, such as electric shocks or apomorphine-induced nausea. It is extremely difficult to assess the efficacy of these methods in human addicts, and laboratory tests in dependent animals offer one means of comparison. We describe here tests of three conceptually different procedures, each reflecting one of the approaches described above. The dependence behaviour was oral self-administration of morphine solutions (morphine HCl 0.5 mg ml⁻¹) by rats which had previously learned to drink such solutions in preference to water⁴.

Sixty male hooded rats aged about 60 d were given solutions of morphine to drink instead of tapwater for 52 d, and from then on their access to fluids was restricted to 2 h each day,

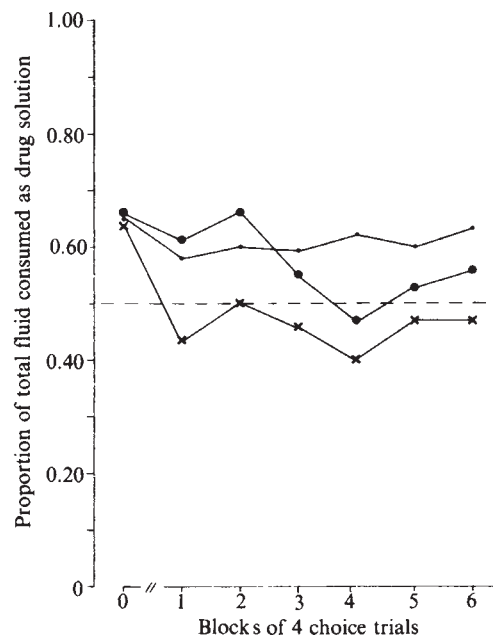


Fig. 1 Injections of morphine (x), naloxone (●) or saline (○) shortly before choice tests do not produce any substantial effects on morphine intake in dependent rats which have learned to drink morphine solutions in preference to water, although during the first 4 d (trial block 1), there is a small reduction in the preferences of the rats injected with morphine. No further adjustment of drug intake occurs in spite of large increases of injected doses of either morphine or naloxone (trial blocks 4-6). Each trial block is the average of four successive choice tests.

between 1100 h and 1300 h. After 14 d of restricted access to only morphine solutions, the rats were given 12 choice tests between morphine solutions and water, alternating with 12 d when morphine solutions alone were provided. Finally, choice tests were given every day for 10 d. The side of the cage where the morphine bottle was presented was varied according to a modified Gellerman series. Drinking was measured by weighing the fluid bottles and correcting for spillage. A normal (unreversed) cycle of lighting was maintained and room temperature averaged $22 \pm 2^\circ \text{C}$. The rats were weighed at regular intervals and their normal pellet diet was available *ad libitum* throughout.

Thirty-nine rats showing high and consistent preferences for morphine were matched on the basis of morphine consumption and were allocated to 'treatment' groups as follows: 'replacement' ($n = 10$), 'blockade' ($n = 10$) and controls ($n = 9$); a fourth group ($n = 10$) was given the aversion procedure. The rats treated by replacement were injected with morphine HCl 2 h before each daily choice trial; the dose was 30 mg per kg body weight for 12 d and was then increased at 4-d intervals to 60, 90 and finally 120 mg per kg. The rats treated by antagonist blockade received daily injections of naloxone HCl 0.5 h before each choice trial; the dose was 1 mg per kg for 12 d, increasing to 2, 4 and then 8 mg per kg over the next 12 d. Controls were injected with normal saline 2 h before choice trials. All injections were given intraperitoneally in a volume of 2 ml per kg.

Figure 1 shows that before treatments were begun (choice trial block 0) the rats were taking about 65% of their daily fluids as morphine solution, corresponding to an absolute dose of 20-25 mg per kg. After the injections were begun (trial blocks 1-6), the control group continued drinking morphine in a virtually unchanged way, and comparisons between the treated and control groups revealed no significant differences overall or at any point (analyses of variance on arc-sin transformed scores of proportional morphine intake). But within the morphine-injected group consumption of morphine fell to 43% during the first 4 d of medication (trial block 1, $t = 3.47$, $P < 0.01$). Increasing the dose of morphine up to six times the

amount usually drunk had no further selective effect on morphine consumption (trial blocks 4–6), although total fluid intake was reduced by 16% ($P = <0.05$), indicating a general depressant effect of the high dose of injected morphine.

Pretreatment with nalorphine⁵ and naloxone⁶ depresses responding for intravenous doses of morphine in rats, and naloxone 'blockade' can reduce preferences for solutions of etonitazine⁷. In our experiment, pretreatment with naloxone (up to 8 mg per kg) produced neither an initial increase in morphine consumption⁸ nor an eventual extinction of preferences. An even larger dose of naloxone, or injections of naltrexone, which is an antagonist with a longer action, might have been more effective, but other measures showed that both naloxone and morphine in the chosen dose ranges were exerting marked effects. The controls grew steadily and gained about 20 g over the 24 d of injections, whereas both the drugged groups failed to gain weight during the first 12 d of constant dosage. Increasing the dose of morphine to 120 mg per kg per d led to more weight loss, but, paradoxically, the rats given more naloxone began to gain weight; however, they remained extremely difficult to handle and had copious diarrhoea.

An aversion procedure was tested in 10 dependent rats (Fig. 2). Rats can develop long-lasting aversions to novel tastes if such tastes are associated with 'sickness', and conditioned aversions to novel flavours paired with the morphine abstinence state^{10,11} have been demonstrated. The morphine-drinking rats were therefore given their usual choice trials, but every third day they were instead given 0.5 h access to a solution with a new taste. This novel solution was the usual morphine 0.5 mg ml⁻¹, but with added saccharin sodium (0.5 mg ml⁻¹). Immediately after access to the morphine plus saccharin solution, the rats were injected with naloxone (1 mg per kg). Thus, the novel taste was paired on four occasions with precipitated abstinence from morphine. During intervening

pairs of days, the rats continued with choice trials between morphine solution and water (choice trials pairs 1–4, Fig. 2). No further aversion training was given, and then, for the next 12 d, the rats were given either a choice between morphine and water as before, or, on alternate days, a choice between morphine plus saccharin and water. They clearly rejected the mixture, while maintaining self-administration of morphine ($t = 8.67$, $P < 0.005$). During the next 12 d (choice trial pairs 11–16), these rats continued to reject the solution of morphine plus saccharin, even though the consequence was 'self-imposed' morphine abstinence. Choice trials pairs 17–22 show that although saccharin alone was initially rejected, possibly reflecting enhanced neophobia¹², the rats very soon began to consume it in increasing amounts while still rejecting the mixture. The fifth phase (choice trial pairs 23–28) demonstrated that even after 24 d of 'voluntary' abstinence, the rats readily relapsed to morphine drinking when given morphine–water choice trials again. The slight reduction in the proportional consumption of morphine was not significant ($t = 1.14$). Tests in other groups of rats that are familiar with the taste of morphine, have confirmed that solutions of morphine plus saccharin are not rejected spontaneously¹³.

In contrast with other studies^{5,8,14}, these experiments provide minimal evidence of regulation of opioid intake by dependent subjects and they emphasise the role of conditioned reinforcers^{6,15,16} in sustaining morphine-seeking behaviour. The high relapse rate of the averted (24-d abstinent) rats also stresses the importance of learned tendencies. The ability of the averted rats to make subtle discriminations between tastes that might or might not be 'safe' underlines a major limitation of behavioural treatment techniques, which is a lack of generalisation of learning to stimuli outside the treatment setting.

This work was supported by the MRC; A. R. Teixeira is a British Council scholar. We thank Endo Laboratories for gifts of naloxone.

L. MUMFORD
A. R. TEIXEIRA
R. KUMAR

Institute of Psychiatry, de Crespigny Park,
London SE5, UK

Received 14 October 1977; accepted 13 January 1978.

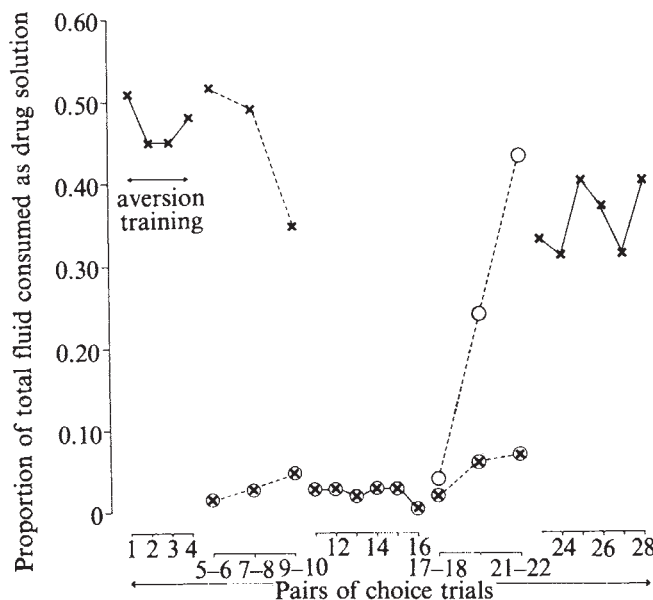


Fig. 2 An aversion to a novel flavour (morphine plus saccharin, x) can be induced in rats which have learned to prefer morphine solutions (o) to water. This aversion is long-lasting (trial pairs 5–22), and dependent rats abstain from morphine if the solutions to which they have been averted are made the only source of the drug. Giving choice tests on alternate days between morphine and water or morphine plus saccharin and water (trial pairs 5–10) shows that the aversion is selective, since the rats continue to ingest plain morphine with which they are familiar. Similar comparisons with saccharin alone (o) (trial pairs 17–22) show that this novel taste is initially rejected, but then the rats drink saccharin in rapidly increasing amounts while continuing to reject morphine plus saccharin. In spite of prolonged abstinence from morphine, the animals relapse when offered plain morphine solutions again (trial pairs 23–28).

- Dole, V. P. & Nyswander, M. E. *J. Am. med. Ass.* 193, 645–650 (1965).
- Martin, W. R. *et al. Arch. gen. Psychiat.* 28, 286–295 (1973).
- Martin, W. R. & Sandqvist, U. L. *Archs gen. Psychiat.* 30, 31–33 (1974).
- Kumar, R., Steinberg, H. & Stolerman, I. P. *Nature* 218, 564–565 (1968).
- Goldberg, S. R., Woods, J. H. & Schuster, C. R. *J. Pharmac. exp. Ther.* 176, 464–471 (1971).
- Davis, W. M. & Smith, S. G. *Biol. Psychiat.* 9, 181–189 (1974).
- Meyer, R. E., Marcus, R., Carnathan, G. & Cochlin, J. *Psychopharmacology* 47, 273–279 (1976).
- McMillan, D. E. *et al. J. Pharmac. exp. Ther.* 196, 269–279 (1976).
- Garcia, J. & Koelling, R. A. *Psychon. Sci.* 4, 123–124 (1966).
- Ternes, J. W. *Bull. psychon. Soc.* 5, 311–312 (1975).
- Pilcher, C. W. T. & Stolerman, I. P. *Pharmac. Biochem. Behav.* 4, 159–163 (1976).
- Best, M. R. & Batson, J. D. *J. exp. Psychol.* 3, 132–143 (1977).
- Kumar, R., Mumford, L. & Teixeira, A. R. *Br. J. Pharmac.* 60, 276P (1977).
- Stolerman, I. P. & Kumar, R. *Psychopharmacologia* 17, 137–150 (1970).
- Kumar, R. & Stolerman, I. P. *J. comp. physiol. Psychol.* 78, 457–465 (1972).
- Wikler, A., Pescor, F. T., Miller, D. & Norell, H. *Psychopharmacologia* 20, 103–117 (1971).

Serotonergic component of neuroleptic receptors

BINDING studies performed *in vitro* with ³H-haloperidol^{1–3}, ³H-spiperone^{4–6}, ³H-dopamine^{1–3,6} and ³H-apomorphine^{7,8} showed that the specific neuroleptic binding sites in the striatum are dopaminergic. The frontal cortex, however, was much more labelled by ³H-spiperone than by ³H-haloperidol^{9,10}. The investigations reported here indicate that the neuroleptic receptor sites in the frontal cortex labelled by ³H-spiperone are predominantly serotonergic and virtually identical to those labelled by ³H-LSD. These conclusions were reached from a study of the relative binding affinities of a series of serotonin or dopamine agonists or antagonists for rat frontal cortex and striatal receptors. Significant correlations were obtained between the binding activities in the rat frontal cortex and *in vivo*

potencies in the pharmacological anti-tryptamine test. Similarly, binding activities in the striatum were significantly correlated with the potencies in the anti-apomorphine test. We therefore suggest that serotonergic, as well as dopaminergic, receptors are involved in the mechanism of action of neuroleptic drugs.

The methods used to assay *in vitro* and *in vivo* receptor binding were essentially similar to those described previously^{4,10,11}. *In vitro* experiments were carried out with a total particulate tissue preparation of rat striatum or frontal cortex. Incubation conditions were the same in all experiments: ³H-ligand concentration of 2 nM; 11.4 mg tissue ml⁻¹; NaCl, 120 mM; KCl 5 mM; CaCl₂ 2 mM; MgCl₂, 1 mM; pargyline 10 μ M; ascorbic acid 0.1% in Tris-HCl buffer 0.05 M, pH 7.6; volume 2.2 ml, 37 °C. After 10 min incubation was stopped by rapid filtration under suction through Whatman GF/B glass fibre filters. These were washed twice with 5 ml ice-cold buffer and then counted for radioactivity. Stereospecific binding of ³H-spiperone and ³H-haloperidol was taken as the difference between radioactivity on the filter in assays in the presence of (–) and (+)-butaclamol. When present in excess (10⁻⁵ M) the pharmacologically active (–)-enantiomer completely displaced the labelled ligand from its specific binding sites, whereas the inactive (+)-enantiomer had no effect. Blank values for ³H-LSD binding were measured in the presence of excess cold LSD (10⁻⁶ M). Inhibition of specific binding by drugs was measured at 5 different concentrations and expressed as IC₅₀ values (that is, drug concentration causing 50% inhibition of the specific binding of the labelled ligand). Male Wistar rats weighing 240 $\pm$ 20 g were used for the *in vivo* pharmacological experiments. A detailed description of the methods has been reported previously¹², but briefly, rats were pretreated with the drugs and then injected intravenously with apomorphine-HCl (1.25 mg kg⁻¹) 0.5 h later. Stereotypy was scored every 5 min over a 1-h period. Thereafter, the animals were injected i.v. with 40 mg kg⁻¹ tryptamine and bilateral clonic seizures of the forepaws were assessed immediately after injection. Based on control results, 'all or none' criteria were selected and used for the calculation of the ED₅₀ values in μ mol kg⁻¹. Multiple regression analysis was carried out using a AQSARF (automatic quantitative structure activity finder) program¹³.

In the rat frontal cortex *in vitro* stereospecific binding of ³H-spiperone amounted to 23 pmol per g tissue. This represented 70% of the total binding and half the amount of stereospecific binding in the striatum. ³H-LSD also showed a high specific binding in the frontal cortex (9.8 pmol per g tissue). For ³H-haloperidol, however, the binding in the frontal cortex was ten times lower than in the striatum⁹.

Specific *in vitro* binding in the frontal cortex of both ³H-spiperone and ³H-LSD was inhibited more by serotonin-like compounds than by dopamine agonists (Table 1). The reverse was true for *in vitro* binding of ³H-haloperidol and ³H-spiperone in the striatum. Serotonin antagonists were more potent inhibitors in binding tests in the frontal cortex than in the striatum (Table 1). Comparison of the binding results with the results of other *in vivo* pharmacological tests (Table 1) shows that binding in the frontal cortex is correlated with the anti-tryptamine activity of the drugs. From Table 1 it can also be seen that the order of compounds measured by the ratio of their *in vitro* binding activity (IC₅₀ in ³H-haloperidol binding in the striatum to the IC₅₀ of ³H-spiperone binding in the frontal cortex) closely parallels that obtained from the ratio of their *in vivo* pharmacological activity (ED₅₀ in anti-tryptamine test to the ED₅₀ in the anti-apomorphine test).

Table 2 shows that in the series of neuroleptics of different chemical classes the IC₅₀-values of *in vitro* binding in the frontal cortex and in the striatum varied widely. Regression analysis of the IC₅₀ values of the *in vitro* binding tests revealed a highly significant correlation between the values obtained for ³H-spiperone and ³H-LSD binding in the frontal cortex (correlation coefficient $r = 0.93$) and also between the values for ³H-spiperone and ³H-haloperidol binding in the striatum ($r = 0.92$). A linear correlation was not obtained between the IC₅₀ values of two different brain areas. Certain compounds were more active in the frontal cortex binding test than in the striatal one. Pipamperone showed the highest dissociation (ratio of *in vitro* IC₅₀ values = 20), having a ratio even higher than the serotonin agonists and most of the serotonin antagonists. On the other hand, the benzamides were 100 to 1,000 times more active in the striatal binding tests but were only weakly active overall. Specific and potent dopamine antagonists, like

Table 1 *In vitro* potency of (a) amine agonists and (b) 5-hydroxytryptamine (serotonin) antagonists in binding tests using rat striatum and frontal cortex and *in vivo* antagonism of apomorphine-induced stereotypy and tryptamine-induced clonic seizures in rats

	Inhibition of specific binding –log IC ₅₀ (M) $\pm$ s.e.m.*				Antagonism <i>in vivo</i> ED ₅₀ (μ mol kg ⁻¹)		Ratio	
	Striatum ³ H-spiperone	³ H-haloperidol	Frontal cortex ³ H-spip.	³ H-LSD	Apomorphine stereotypy	Tryptamine seizures	<i>In vitro</i> † activities	<i>In vivo</i> ‡ activities
a								
Bufotenin	4.40 $\pm$ 0.05	4.7 $\pm$ 0.1	5.8 $\pm$ 0.1	6.20 $\pm$ 0.07			13	
5-Hydroxytryptamine	4.10 $\pm$ 0.05	4.5 $\pm$ 0.1	5.5 $\pm$ 0.1	6.1 $\pm$ 0.1			10	
LSD	7.25 $\pm$ 0.05	7.3 $\pm$ 0.1	7.6 $\pm$ 0.1	8.3 $\pm$ 0.1			2	
Tryptamine	4.05 $\pm$ 0.05	4.6 $\pm$ 0.1	4.78 $\pm$ 0.07	5.3 $\pm$ 0.1			1.5	
Apomorphine	6.3 $\pm$ 0.1	7.3 $\pm$ 0.1	5.0 $\pm$ 0.1	5.5 $\pm$ 0.1			0.005	
Dopamine	5.2 $\pm$ 0.1	6.0 $\pm$ 0.1	3.6 $\pm$ 0.1	3.7 $\pm$ 0.1			0.004	
2-(N, N-dipropyl) amino- 5, 6-dihydroxy tetralin	6.9 $\pm$ 0.2	7.8 $\pm$ 0.1	4.1 $\pm$ 0.1	4.4 $\pm$ 0.1			0.0002	
b								
Mianserine	5.7 $\pm$ 0.1	5.8 $\pm$ 0.1	7.4 $\pm$ 0.1	7.35 $\pm$ 0.05	266	2	40	133
Cinanserine	5.0 $\pm$ 0.1	5.4 $\pm$ 0.1	6.90 $\pm$ 0.05	6.95 $\pm$ 0.05	452	9.5	32	48
Methysergide	6.5 $\pm$ 0.1	6.3 $\pm$ 0.1	7.45 $\pm$ 0.05	8.10 $\pm$ 0.05	120	1.5	14	80
Pizotifen	6.3 $\pm$ 0.2	6.6 $\pm$ 0.1	7.7 $\pm$ 0.1	7.7 $\pm$ 0.1	102	0.9	13	113
Cyproheptadine	6.7 $\pm$ 0.1	7.1 $\pm$ 0.1	7.7 $\pm$ 0.1	7.55 $\pm$ 0.05	65	1	4	65

*Mean values were calculated from three independent determinations using freshly prepared tissue and drug solutions, and performed in duplicate.

Presently shown IC₅₀ values were not converted into K_i values using the Cheng-Prusov equation, since both receptor populations in striatum and frontal cortex seemed to be heterogeneous and the real dissociation constant of the labelled ligand cannot be measured.

†Ratio of *in vitro* activities is calculated as IC₅₀ (halo., striatum)/IC₅₀ (spip., frontal cortex), expressing within the series for the various compounds differences in dissociation between affinity for frontal cortical and striatal receptors.

‡Ratio of *in vivo* activities is calculated as ED₅₀ in apomorphine test to ED₅₀ in tryptamine test.

Table 2 *In vitro* potency of 23 neuroleptics in binding tests using rat striatum and frontal cortex and *in vivo* antagonism of apomorphine-induced stereotypy and tryptamine-induced clonic seizures in rats

	Inhibition of specific binding —log IC ₅₀ (M) ± s.e.m.*				Antagonism <i>in vivo</i> ED ₅₀ (μmol kg ⁻¹)		Ratio	
	Striatum ³ H-spi. (A)	³ H-halo. (B)	Frontal cortex ³ H-spi. (C)	³ H-LSD (D)	Apomorphine stereotypy (F)	Tryptamine seizures (F)	<i>In vitro</i> † activities	<i>In vivo</i> activities E/F
Pipamperone	5.9 ± 0.1	6.5 ± 0.2	7.79 ± 0.06	7.57 ± 0.07	14.3	1.55	19.5	9.2
Clozapine	5.9 ± 0.2	6.4 ± 0.2	7.32 ± 0.09	7.08 ± 0.07	32.7	9.39	8.3	3.5
Azaperone	6.50 ± 0.05	6.7 ± 0.2	7.45 ± 0.06	7.5 ± 0.1	2.35	1.35	5.6	1.7
Chlorprothixene	7.25 ± 0.05	7.55 ± 0.05	8.00 ± 0.05	7.61 ± 0.09	1.08	1.45	2.8	0.7
Clothiapine	7.1 ± 0.1	7.4 ± 0.2	7.74 ± 0.08	7.4 ± 0.1	0.38	1.11	2.2	0.3
Chlorpromazine	6.85 ± 0.07	6.9 ± 0.1	7.21 ± 0.07	6.6 ± 0.1	0.82	2.48	2.0	0.3
Methitapine	8.0 ± 0.2	8.0 ± 0.1	8.23 ± 0.07	7.95 ± 0.05	0.20	0.20	1.7	1.0
Propericiazine	7.90 ± 0.05	8.0 ± 0.1	8.1 ± 0.1	7.2 ± 0.1	0.23	0.90	1.3	0.3
Flupenthixol	7.23 ± 0.05	7.7 ± 0.2	7.40 ± 0.05	6.75 ± 0.05	0.26	0.82	0.5	0.3
Thioridazine	6.65 ± 0.05	7.4 ± 0.2	6.96 ± 0.07	6.8 ± 0.1	13.1	26.3	0.4	0.5
Fluspiperone	8.74 ± 0.04	9.4 ± 0.1	8.68 ± 0.05	8.1 ± 0.1	0.032	0.12	0.2	0.3
Fluphenazine	7.41 ± 0.07	7.8 ± 0.2	7.0 ± 0.1	6.6 ± 0.1	0.13	1.14	0.2	0.1
Droperidol	8.1 ± 0.1	8.7 ± 0.1	7.90 ± 0.05	7.57 ± 0.07	0.035	1.29	0.2	0.03
Spiperone	8.90 ± 0.03	9.4 ± 0.1	8.45 ± 0.05	8.2 ± 0.1	0.053	0.21	0.1	0.2
(+)-Butaclamol	7.7 ± 0.1	8.0 ± 0.2	7.0 ± 0.1	7.38 ± 0.06	0.28	1.68	0.1	0.2
Trifluoperidol	7.85 ± 0.06	8.8 ± 0.1	7.55 ± 0.05	6.90 ± 0.05	0.040	0.74	0.06	0.05
Benperidol	8.35 ± 0.06	9.05 ± 0.05	7.70 ± 0.05	7.6 ± 0.1	0.023	0.76	0.04	0.03
Pimozide	7.80 ± 0.05	8.5 ± 0.1	7.0 ± 0.1	6.7 ± 0.1	0.10	> 22	0.03	0.004
Clopimozide	7.03 ± 0.06	8.1 ± 0.3	6.5 ± 0.1	6.45 ± 0.05	0.30	> 80	0.03	0.004
Haloperidol	7.69 ± 0.06	8.5 ± 0.1	6.83 ± 0.07	6.15 ± 0.05	0.043	2.05	0.02	0.02
Oxipromide	7.7 ± 0.4	7.93 ± 0.08	6.05 ± 0.05	6.35 ± 0.05	0.11	7.95	0.01	0.01
Metoclopramide	5.8 ± 0.1	6.95 ± 0.05	4.93 ± 0.08	5.15 ± 0.05	1.69	41.4	0.01	0.04
Sulpiride	6.06 ± 0.06	7.1 ± 0.1	4.1 ± 0.1	< 5	62.8	938	0.001	0.07

In this series of compounds statistically significant correlations ($P < 0.01$) were obtained between the activities of A and B ($r = 0.92$), C and D ($r = 0.93$), A and E ($r = 0.81$), A and E + F ($r = 0.84$), B and E ($r = 0.85$), C and F ($r = 0.83$), D and F ($r = 0.74$).

*Mean values were calculated from three independent determinations using freshly prepared tissue and drug solutions, and performed in duplicate.

Presently shown IC₅₀ values were not converted into K_i values using the Cheng-Prusov equation, since both receptor populations in striatum and frontal cortex seemed to be heterogeneous and the real dissociation constant of the labelled ligand cannot be measured.

†Ratio of *in vitro* activities is calculated as IC₅₀ (halo., striatum)/IC₅₀ (spi., frontal cortex), expressing within the series for the various compounds differences in dissociation between affinity for frontal cortical and striatal receptors.

pimozide and clopimozide^{14,15}, were 30 to 40 times more active inhibitors of haloperidol binding in the striatum than of spiperone binding in the frontal cortex.

Regression analysis (correlation coefficients in Table 2) established that *in vitro* binding potency in the frontal cortex was highly correlated with *in vivo* potency in the antagonism of bilateral clonic seizures induced by tryptamine. The IC₅₀ values of the *in vitro* striatal binding tests were correlated with *in vivo* antagonism of apomorphine stereotypy. Also noteworthy is the highly significant correlation shown by means of the two-variable regression equation which relates spiperone

binding in the striatum to anti-apomorphine and anti-tryptamine activity simultaneously [$-\log \text{IC}_{50} (A) = -0.54 \log \text{ED}_{50} (E) - 0.33 \log \text{FD}_{50} (F) + 7.2$] ($r = 0.84$, $P < 0.01$) (see Table 2 for significance of A, E, F). A similar multivariable correlation could not be obtained between any of the other binding data.

The *in vitro* finding that neuroleptics bind to both striatal and frontal cortex receptors, and that certain compounds are more selective for one brain area was further confirmed by *in vivo* binding studies (Table 3). *In vivo* displacement of ³H-spiperone from its receptors by pipamperone and LSD was readily obtained at low doses in the frontal cortex, but not in the brain

Table 3 *In vivo* displacement of labelled spiperone in rat brain

Region	Control	Labelled spiperone (pg mg ⁻¹)				Haloperidol 2.5 mg kg ⁻¹
		LSD 0.63 mg kg ⁻¹	10 mg kg ⁻¹	Pipamperone 0.63 mg kg ⁻¹	10 mg kg ⁻¹	
After 2 h						
Striatum	3.8±0.1	3.7±0.1	3.2±0.1†	4.0±0.3	3.6±0.1	2.9±0.2‡
Nucleus accumbens	3.6±0.2	3.5±0.3	3.2±0.1	3.5±0.3	3.2±0.2	2.7±0.0*
Tuberculum olfactorium	2.5±0.1	2.5±0.2	2.0±0.2*	2.1±0.3	2.2±0.2	2.2±0.1*
Frontal cortex	1.5±0.04	0.8±0.1‡	0.65±0.02‡	0.94±0.02‡	0.65±0.02‡	0.83±0.02‡
Cerebellum	0.40±0.03	0.37±0.01	0.40±0.01	0.38±0.02	0.36±0.02	0.35±0.01
After 4 h						
Striatum	3.7±0.1	4.3±0.3	3.8±0.1	3.8±0.1	2.7±0.1‡	2.1±0.1‡
Nucleus accumbens	3.4±0.2	3.9±0.3	3.0±0.1	3.2±0.1	2.4±0.3*	2.1±0.1‡
Tuberculum olfactorium	2.2±0.1	2.1±0.2	1.9±0.1	2.2±0.2	1.5±0.1‡	1.4±0.1‡
Frontal cortex	0.8±0.1	0.48±0.02‡	0.43±0.03‡	0.62±0.06*	0.37±0.01‡	0.44±0.01‡
Cerebellum	0.27±0.02	0.22±0.01	0.25±0.01	0.30±0.03	0.23±0.01	0.27±0.01

Rats were given ³H-spiperone (0.005 mg kg⁻¹, specific activity 9 Ci mmol⁻¹) intravenously. One hour later, unlabelled LSD, pipamperone and haloperidol were similarly injected. Two and four hours after the labelled drug, radioactivity was determined in various brain regions. Each value is the mean of four determinations ± s.e.m. Significant differences (by Student's *t* test) from control values are indicated

* $P < 0.05$ † $P < 0.01$ ‡ $P < 0.001$.

dopaminergic areas. Radioactivity in dopaminergic areas was reduced by high doses of LSD, but only significantly after 2 h. In contrast, haloperidol supplanted ^3H -spiperone in both dopaminergic areas and in the frontal cortex.

These results show that spiperone and LSD label the same receptors in the frontal cortex and that these receptors are different from those labelled by spiperone and haloperidol in the striatum. In *in vitro* as well as *in vivo* binding experiments, neuroleptics showed high affinity for the receptors in both brain areas. Different observations provided evidence that receptors in the frontal cortex are predominantly of serotonergic nature: the receptors could be differentiated by serotonin agonists and antagonists; LSD has previously been described as a ligand for serotonergic receptors¹⁶; and a direct correlation has been shown between relative affinities for frontal cortical receptors and *in vivo* antagonism of the tryptamine test. Tryptamine is assumed to mimic serotonin action at central and peripheral receptor sites^{17,18}.

In confirmation of previous reports, striatal receptors seemed to be chiefly dopaminergic, as evidenced by the correlation between striatal *in vitro* binding activity and the *in vivo* potency in the apomorphine test. It has frequently been reported that apomorphine mimics central dopaminergic effects (see refs in ref. 19). At present there is little evidence of the homogeneity or heterogeneity of the receptor populations in the brain areas studied. Kinetic investigations have indicated the presence of different binding sites in the striatum⁴, and these were more apparent with ^3H -spiperone than with ^3H -haloperidol. The two-variable regression equation presented here (between A and $E+F$) also points to a minor serotonergic component in the striatal binding sites. This is supported by the findings of Withaker and Seeman²⁰ that hallucinogens bind to neuroleptic receptors in the calf caudate. The serotonergic component, however, did not become apparent in the binding of ^3H -haloperidol in the striatum, therefore this test was chosen more as a pure *in vitro* assessment of the anti-dopaminergic activity of drugs. Investigations of the kinetics of ^3H -spiperone binding in the frontal cortex also pointed to the presence of various different binding sites (our work, in preparation).

The experiments reported here are the first *in vitro* demonstration that both serotonergic and dopaminergic receptors may be involved in the mechanism of action of neuroleptics. The degree of affinity and the ratio of affinity of the drugs for receptors is closely reflected in their pharmacological potency and profile. The precise significance of the anti-serotonergic component of neuroleptics, however, requires assessment of its clinical implications. Note that pipamperone, the compound with the highest dissociation in binding affinity in favour of the frontal cortex, is noted for its particular properties in clinical psychiatry. The drug is known for its anti-agitation effects²¹⁻²⁴ and its conspicuous capacity to normalise disturbed sleep rhythms in psychiatric patients²⁵. Pimozide, however, which had very pronounced striatal activity, is a potent anti-psychotic with a profile which is different from other neuroleptics²⁶. The present observations do not detract from the dopamine theory on schizophrenia, but they do imply that a broadening of this theory to cover serotonergic processes may increase our understanding of this complexity of disorders.

We thank Dr Janssen for constructive criticisms and R. Geerts and W. Gommeren for technical assistance. Part of this work was supported by a grant from IWONL.

JOSÉE E. LEYSEN
CARLOS J. E. NIEMEGEERS
JAN P. TOLLENAERE
PIERRE M. LADURON

Department of Pharmacology,
Janssen Pharmaceutica Research Laboratories,
B-2340 Beerse, Belgium

1. Seeman, P., Chau-Wong, M., Tedesco, J. & Wong, K. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4376-4380 (1975).
2. Creese, I., Burt, D. R. & Snyder, S. H. *Life Sci.* **17**, 993-1002 (1975).
3. Burt, D. R., Enna, S., Creese, I. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4655-4659 (1975).
4. Leysen, J. E., Gommeren, W. & Laduron, P. M. *Biochem. Pharmac.* **27**, 307-316 (1978).
5. Laduron, P. M., Janssen, P. F. M. & Leysen, J. E. *Biochem. Pharmac.* **27**, 317-321 (1978).
6. Enna, S. J., Bennett, J. P., Burt, D. R., Creese, I. & Snyder, S. H. *Nature* **263**, 339-341 (1976).
7. Seeman, P., Lee, T., Chau-Wong, M. & Wong, K. *Nature* **261**, 717-719 (1976).
8. Seeman, P., Lee, T., Chau-Wong, M. & Wong, K. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4354-4358 (1976).
9. Leysen, J. & Laduron, P. *Life Sci.* **20**, 281-288 (1977).
10. Laduron, P. & Leysen, J. *Biochem. Pharmac.* **26**, 1003-1007 (1977).
11. Leysen, J., Tollenaere, J. P., Koch, M. H. J. & Laduron, P. *Eur. J. Pharmac.* **43**, 253-267 (1977).
12. Niemegeers, C. J. E., Lenaerts, F. M., Artois, K. S. K. & Janssen, P. A. J. *Archs int. Pharmacodyn.* **227**, 238-253 (1977).
13. Tollenaere, J. P., Moereels, H. & Protiva, M. *Eur. J. med. Chem.* **11**, 293 (1976).
14. Janssen, P. A. J. *et al. Arzneim.-Forsch.* **18**, 261 (1968).
15. Janssen, P. A. J., Niemegeers, C. J. E., Schellekens, K. H. L., Lenaerts, F. M. & Wauquier, A. *Arzneim.-Forsch.* **25**, 1287 (1975).
16. Bennett, J. P., Jr. & Snyder, S. H. *Molec. Pharmac.* **12**, 373-389 (1976).
17. Tedeschi, D. M., Tedeschi, R. E. & Fellows, E. J. *J. Pharmac. exp. Ther.* **126**, 223-232 (1959).
18. Anden, N. E., Corrodi, H. & Fuxe, K. *J. Pharmac. exp. Ther.* **179**, 236-249 (1971).
19. Colpaert, F. C., Van Bever, W. F. M. & Leysen, J. E. M. F. *Int. Rev. Neurobiol.* **19**, 225-267 (1976).
20. Whitaker, P. M. & Seeman, P. *J. Pharm. Pharmac.* **29**, 506-507 (1977).
21. Van Hemert, J. *Acta psychiatr. scand.* **52**, 237-245 (1975).
22. De Bort, R. *Acta psychiatr. belg.* **76**, 157-166 (1976).
23. Van Rensynghe de Votvrie, G. & De Bie, M. *Acta psychiatr. belg.* **76**, 688-695 (1976).
24. Squelart, P. & Saravia, J. *Acta psychiatr. belg.* **77**, 284-293 (1977).
25. Asoms, C., De Backer-Dierick, G. & Vereecken, J. L. P. M. *Acta psychiatr. scand.* **55**, 116-122 (1977).
26. Wouters, A. C. R. *Congrès Psychiatrie. Neurol. de langue Française, Session 67, Bruxelles, 1969*, 561-566.

Functions of calcium in sweat secretion

CALCIUM is an important regulatory factor in many physiological systems, including fluid-transporting epithelia. There is good evidence that such regulatory factors are involved in normal secretory function¹, and indirect evidence that they may be involved in secretory irregularities in disease states^{2,3}. Using a novel *in vitro* preparation, we have examined the effects of calcium on the secretory properties of the human eccrine sweat gland. The sweat gland consists of a single coiled tubule which is closed at the proximal secretory end and open at the distal re-absorptive end. On stimulation, the secretory portion of the tubule secretes an isotonic precursor fluid from which NaCl is absorbed hypertonically in the distal, re-absorptive duct. The final secretory product, hypotonic sweat, is excreted on to the epidermis for evaporation. We have found that calcium is an essential requirement for stimulating sweat and that this ion may be a factor in regulating the hypotonicity of sweat.

Skin biopsies were taken from the upper back of 21-39-yr-old male volunteers with an electric biopsy punch 3 mm in diameter. The skin plugs were mounted in 15-ml chambers such that the epidermis was continuously immersed in hexadecane while the dermal and subdermal tissues were bathed in oxygenated Ringer's solution (132 mM NaCl, 5 mM KCl, 1.2 mM MgSO₄, 11 mM glucose, 5 mM NaHEPES and CaCl₂ to the desired concentration; pH 7.4). With this arrangement, small droplets of sweat formed under the oil on the epidermis when the sweat glands were stimulated pharmacologically by adding 5×10^{-7} M mecholyl to the bath. The droplets from active glands (one to three per biopsy plug) were collected in constant bore capillaries between oil blocks. Secretory rates of single glands were measured by dividing the sample volume by the time in minutes during which the sample was secreted (usually 2-5 min). The concentrations of sodium and calcium were determined by microdroplet X-ray analysis⁴. The preparation will be described in detail elsewhere.

As shown in Table 1, secretory rates at a bath calcium concentration of 5.0 and 10.0 mM were significantly lower ($P < 0.05$) than at 2.0 mM, suggesting that at higher extracellular concentrations, calcium may be somewhat inhibitory to active secretion or that it may be acting to decrease the epithelial permeability such that the passive movements of fluid and electrolytes are restricted⁵.

However, secretory activity was found to be entirely dependent on the presence of external calcium in the bath (Table 1). Sweat

secretion could not be stimulated when calcium was omitted from the bath, but the effect was reversible, and a threshold of about 0.1 mM calcium was demonstrated by adding calcium to the bath in 0.05 mM increments. Since secretion was induced by direct stimulation of the receptor site with the cholinergic agent, mecholyl (*O*-acetyl- β -methylcholine chloride), the inhibitory effect is unlikely to be due to presynaptic effects. On the other hand, the findings in salivary glands^{6,7} and pancreas⁸ that membrane depolarisation and intracellular potassium release from secretory cells continue in the absence of calcium, argues that calcium does not mediate receptor binding of the neurotransmitter or membrane phenomena related to depolarisation. Thus, external calcium may be required to elevate intracellular calcium during stimulation to levels necessary for activating and maintaining secretory metabolism⁹.

Calcium also seems to be involved in regulating sodium re-absorption in the duct. In both *in vivo* and *in vitro* measurements at a bath concentration of 2.0 mM, the mean concentration of calcium in collected sweat is about 0.5 mM. However, the actual concentrations for single glands *in vitro* ranged from 0.2 to 0.8 mM calcium. The source of this variation is not understood, but we have observed a positive correlation between calcium and sodium concentrations in collected sweat, even though the bath calcium and sodium remained constant. To examine the appearance of calcium in the sweat and the possibility that calcium might be inducing an inhibitory effect on sodium transport in the re-absorptive duct, we attempted to extend the range of sweat calcium by changing the bath calcium concentration.

Table 1 shows that calcium concentration in the final sweat increased, but not linearly, with external bath calcium concentration. In another set of experiments we observed that in

Table 1 Sweat rates and calcium concentrations for corresponding concentrations of calcium in the bath

Bath [Ca]	0.0	2.0	5.0	10.0
Secretory rates	0	$3.94 \pm 0.29^*$	2.63 ± 0.32	2.41 ± 0.14
Sweat [Ca]	—	0.52 ± 0.05	1.24 ± 0.10	1.60 ± 0.25

Ten biopsied skin plugs were rinsed and incubated for 30 min in Ringer's solution containing 0, 2, 5 or 10 mM calcium before stimulation with mecholyl. Three collections were then taken from each gland when possible. The order of exposure to the different calcium concentrations was randomised to avoid time related effects. Data was taken from a total of 22 glands.

*s.e.m.

collections from three glands bathed in 25 mM calcium, the average final sweat calcium was only 1.4 mM. Unless there is a high capacity for calcium re-absorption in the sweat gland duct, the data indicate that the entry of calcium into the primary sweat is markedly limited at higher calcium concentrations, possibly suggesting a saturable carrier. On the other hand, the possibility that the bath does not serve as the immediate source of sweat calcium cannot be excluded at this point. Although the mechanisms regulating the appearance of calcium in the sweat remain unclear, it is clear that increases in calcium up to about 1.5 mM in the re-absorptive duct markedly inhibit sodium re-absorption (Fig. 1). As this same sodium dependence on sweat calcium concentration was found within the narrower ranges of sweat calcium occurring at each constant calcium bath, the effects are less likely to be due to increased extracellular calcium than to luminal calcium in the duct.

Inhibitory effects of calcium on sodium transport have been noted in other tissues^{10,11} and interpreted as being due to a calcium-induced decrease in the sodium permeability across the apical membrane of the re-absorptive cell¹⁰. This lower permeability may reduce the passive flux of sodium into the cell and diminish sodium transport by limiting the amount of cytoplasmic sodium available for active extrusion by ($\text{Na}^+ + \text{K}^+$)-ATPase at the basal lateral membrane¹².

In conclusion, calcium is functionally significant in sweat secretion on at least two levels. First, it is an essential requirement for cholinergic stimulus-secretion coupling. Secretion cannot be induced unless it is present at a concentration of at least 0.1 mM.

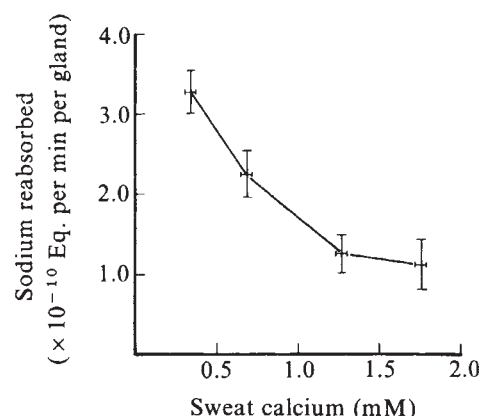


Fig. 1 Sodium re-absorption rate as a function of sweat calcium concentration. The rate of sodium re-absorption is markedly sensitive to concentration of calcium in the final sweat below about 1.2 mM. The rate of sodium re-absorption was calculated from the equation $\text{Na}_r = (137 \text{ mM} - [\text{Na}]_s) \times R$, in which 137 mM is assumed to be the concentration of sodium in the secreted precursor fluid, Na_r is the rate of sodium re-absorption, $[\text{Na}]_s$ is the measured concentration of sodium in the collected sweat, and R is the corresponding sweat rate. Horizontal bars indicate s.e.m. of the sweat calcium concentrations between 0 and 0.5, 0.5 and 1.0, 1.0 and 1.5, and 1.5 and 2.0 mM. Vertical bars represent s.e.m. of the corresponding rates of sodium re-absorption for points within these respective concentration ranges.

Second, the concentration of calcium in the final sweat is inversely correlated with sodium re-absorption. This relationship, together with the indication that the level of sweat calcium may be determined by a secretory component, indicates the presence of a mechanism for regulating the composition of sweat through calcium.

Perhaps it should be noted that certain fluid movements in the disease cystic fibrosis are characterised by an apparent inhibition of sodium transport, and calcium levels in these fluids tend to be higher than normal². These considerations suggest that abnormally high calcium levels in cystic fibrosis secretions may be exerting pathological effects on fluid formation in this disease.

This work was supported by grants from the National Cystic Fibrosis Foundation, the Southern California Kidney Foundation and the NIH (No. 1 R01 AM 20356-01 GMA). P.M.Q. is the recipient of a research career development award (USPHS AM 324-01A1).

C. A. PROMPT
P. M. QUINTON*

Departments of Physiology and Medicine,
University of California Medical School,
Los Angeles, California 90024

Received 7 October; accepted 28 December 1977.

*To whom correspondence should be addressed.

1. Thorn, N. A. in *Secretory Mechanisms of Exocrine Glands* (eds Thorn, N. A. & Petersen, O. H.) (Munksgaard, Copenhagen, 1974).
2. Dearborn, Dorr G. in *Cystic Fibrosis: Projections into the Future* (eds Mangos, J. A. & Talamo, R. C.) (Symposia Specialists, New York, 1976).
3. Prompt, C. A., Quinton, P. M. & Kleeman, C. R. *Nephron* **20**, 4-9 (1977).
4. Quinton, P. M. *Micron* (in the press).
5. Wright, E. M. & Diamond, J. M. *Biochim. biophys. Acta* **163**, 57-74 (1967).
6. Pedersen, O. H., Poulsen, J. H. & Thorn, N. A. *Acta Physiol. Scand.* **71**, 203-210 (1967).
7. Putney, J. W. J. *Physiol., Lond.* **268**, 139-149 (1977).
8. Poulsen, J. H. & Williams, J. A. J. *Physiol., Lond.* **264**, 323-339 (1977).
9. Peterson, O. H. & Ueda, N. J. *Physiol., Lond.* **254**, 583-606 (1976).
10. Curran, P. F. & Gill, J. R. Jr *J. gen. Physiol.* **45**, 625-641 (1962).
11. Schneyer, L. H. in *Secretory Mechanisms of Exocrine Glands* (eds Thorn, N. A. & Petersen, O. H.) (Munksgaard, Copenhagen, 1974).
12. Quinton, P. M. & Tormey, J. McD. *J. Membrane Biol.* **29**, 383-399 (1976).

Allylisopropylacetamide restricts expression of β minor globin gene in Friend cells

Two structural genes, β major and β minor, are present at the β -globin locus (*Hbb*) of a variety of mice phenotypically characterised by adult haemoglobin of the diffuse type¹. These genes may be the result of duplication of an ancestral gene², but there is increasing evidence that their expression is independently regulated. For example, in erythrocytes of the adult

DBA/2 mouse, the ratio of the two nonallelic variant gene products is 4:1 in favour of the β -major chain whereas in Friend erythroleukaemia cells, clone 745, derived from the same mouse strain, the ratio varies from 3 to 0.6, depending on the inducer of erythroid differentiation used³⁻⁶. We reported previously that treatment of clone 745 cells for 6 d with 1×10^{-4} M haemin, an inducer of haemoglobin synthesis⁷, induces synthesis of the

β minor, but not the β major globin chain⁵. This observation suggested the possibility that intracellular levels of haem (the reduced form of haemin) might act physiologically to selectively control the expression of one of the two β globin genes. To test this hypothesis, we determined the effect of allylisopropylacetamide (AIA) on the relative amounts of β globin chains present in induced erythroleukaemia cells. AIA is a strong

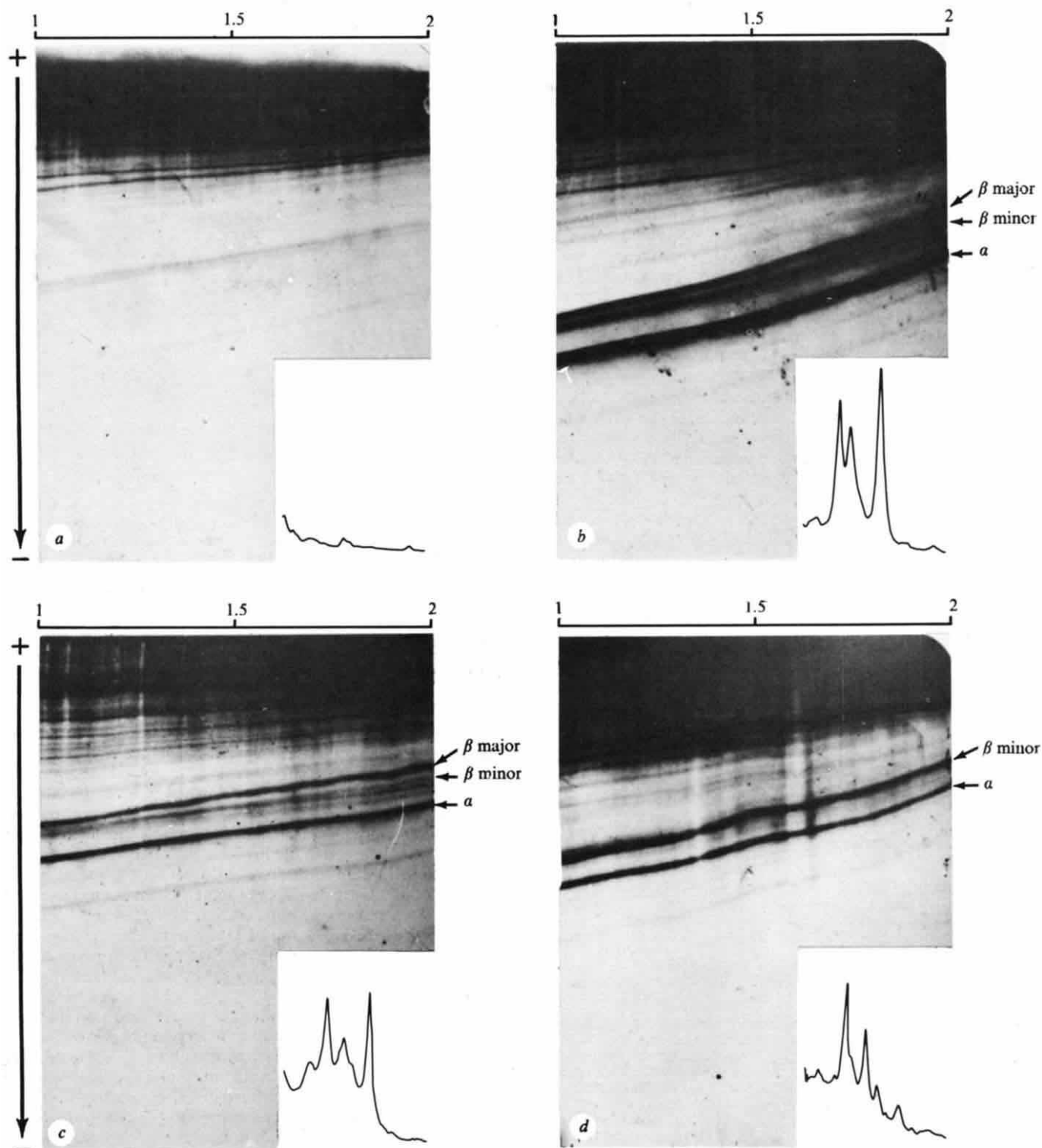


Fig. 1 Fluorographic images of cytoplasmic lysates of Friend cells (clone 745). *a*, Nontreated; *b*, *c*, and *d*, treated for 6 d with: 1.8% DMSO; 1.8% DMSO and $700 \mu\text{g ml}^{-1}$ of AIA; and 1×10^{-4} M haemin, respectively. Insets show scans of the autoradiograms. The ratio between β major and β minor globin chains calculated from them is 1.28 for DMSO-treated and 3.44 for DMSO- and AIA-treated cells. Cytoplasmic lysates (specific activity 200–300 c.p.m. per μg protein) were prepared as described by Kabat *et al.*³ and stored at -20°C . Globin chains were separated by electrophoresis of the lysates in acidic urea gel slabs containing a horizontal gradient (from 0 to 2%) of Triton X-100 and identified as described elsewhere^{5,16}. Fluorograms^{17,18} obtained by exposure of preflashed Kodak RP-omat X-ray film for 3–7 d at -70°C were scanned with a Corning model 750 scanning densitometer equipped with integrating capabilities.

Table 1 Effect of AIA treatment on erythroleukaemia cells

Treatment	Haem concentration (ng per 5×10^6 cells)	% B+ cells	Cell density ($\times 10^6$ per ml)
None	16.8	0	1.92
Induced	135.4	81	1.52
+200 $\mu\text{g ml}^{-1}$ AIA	83.5	78	2.11
+500 $\mu\text{g ml}^{-1}$ AIA	75.6	51	1.95
+700 $\mu\text{g ml}^{-1}$ AIA	59.7	30	1.35

Friend cells, clone 745, were untreated or induced to differentiate with 1.8% DMSO with or without AIA. After 6 d, cell density, percentage of B+ cells and haem concentrations were determined. Intracellular haem concentration was determined by the method described by Sassa²⁰, except that phenol red was not omitted from the culture medium but the cellular pellet was washed three times with phosphate-buffered saline. Values represent averages of duplicates from a representative experiment.

inducer of δ -aminolevulinic acid synthetase in rat hepatocytes⁸ and in mouse erythroleukaemia cells⁹. Induction of the enzyme is due to breakdown by the drug of the haem moiety of the haemoproteins¹⁰. We report here that AIA treatment of Friend cells induced to differentiate by dimethylsulphoxide (DMSO) causes a reduction in the intracellular haem concentration and at the same time a marked decrease in the amount of the β minor globin chain. The ratio of the two β chains thus becomes very close to the ratio observed in normal erythrocytes of adult DBA/2 mouse.

Murine erythroleukaemia cells, clone 745, of DBA/2 origin¹¹ and the culture conditions used in our laboratory have been described elsewhere^{12,13}. For the experiments reported here, cells were induced to differentiate by incubation for 6 d with DMSO 1.8% (w/v). Cells were re-fed for the last 48 h with medium containing 1 $\mu\text{Ci ml}^{-1}$ of a ¹⁴C-amino acid mixture (Amersham, Searle). AIA was added to cultures in concentrations ranging from 200 to 700 $\mu\text{g ml}^{-1}$ at the beginning of the

experiment⁹. At the time of cell collection, the percentage of cells containing haemoglobin [benzidine-positive (B+) cells] was determined by the method of Orkin *et al.*¹⁴.

Mouse globin antiserum was prepared by injecting New Zealand rabbits with haemoglobins of the diffuse type purified by isoelectric focusing¹⁵; its specificity was determined by immunodiffusion tests.

As shown in Table 1, treatment of DMSO-induced Friend cells with increasing concentrations of AIA caused a progressive decrease in the amount of intracellular haem and reduced the percentage of B+ cells. In agreement with the observation of Ebert and Ikawa⁹, no toxicity was observed within the dose range used.

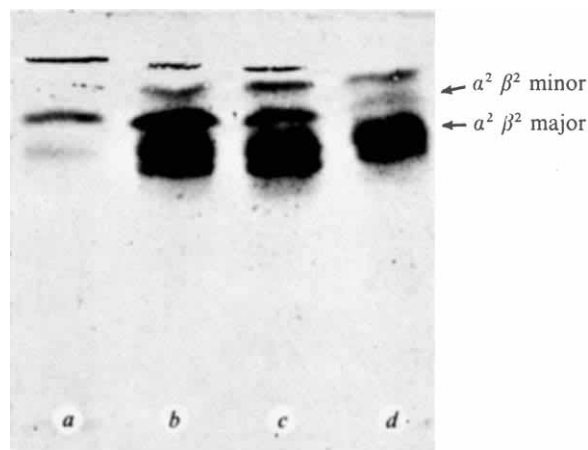


Fig. 3 Isoelectrofocusing separation of haemoglobins present in Friend erythroleukaemia cells. *a*, Haemin-treated (1×10^{-4} M); *b*, DMSO-treated; *c*, hypoxanthine-treated Friend cells. Adult DBA/2 mouse haemoglobins are shown in panel (*d*). Haemoglobins were separated on acrylamide gel slabs as described^{15,18} and the gels were stained with benzidine. The chain composition of each haemoglobin band in control DBA/2 mice was determined as described in detail in ref. 16. Hypoxanthine (6 mM) induced cells were used as an additional control, since we have reported⁵ that hypoxanthine induced β major and β minor globin chains with a ratio similar to that of DMSO-treated cells.

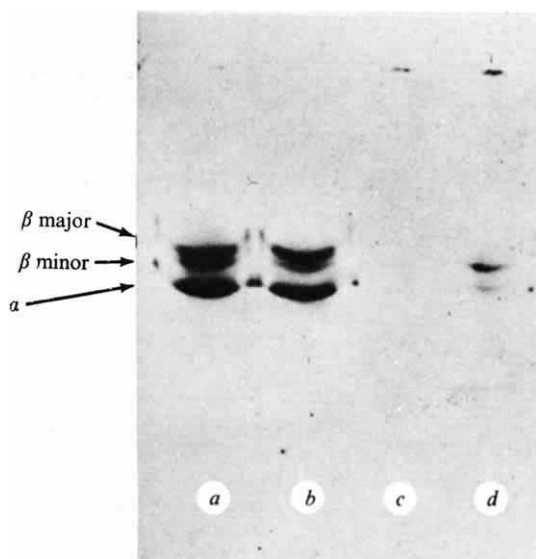


Fig. 2 Differences in globin patterns of immunoprecipitates of cytoplasmic lysates of Friend cells. *a*, DMSO-treated; *b*, DMSO- and AIA- (700 $\mu\text{g ml}^{-1}$) treated; *c*, untreated controls; and *d*, 1×10^{-4} M haemin-treated cells. Cytoplasmic lysates were freed of material that binds nonspecifically to normal rabbit serum and to protein A-bearing *Staphylococcus aureus* Cowan I by two 1-h incubations with excess bacteria (0 °C). Immune complexes were formed by incubating extracts with 100 μl of antiserum for 1 h at 0 °C. The immune complexes were absorbed onto *S. aureus* Cowan I and washed according to the procedure of Cullen and Schwartz¹⁹. The adsorbed proteins were removed from the bacterial surface by resuspension of the complex in 4 M urea, 5% acetic acid followed by centrifugation. The supernatant was subjected to electrophoresis in gel slabs containing 2% Triton as described in Figure 1 legend. The ratios between β major and β minor globin chains calculated from them was 1.35 for DMSO-treated and 3.35 for DMSO- and AIA-treated cells.

The electrophoretic patterns of cytoplasmic lysates of Friend cells induced to synthesise haemoglobin in the presence or in the absence of AIA are shown in Fig. 1. For comparison, the electrophoretic patterns of cytoplasmic lysates of haemin-induced and untreated Friend cells are also shown. The insets in Fig. 1 contain scans of the autoradiograms showing only the portions to which globin chains have migrated. In cells treated for 6 d with DMSO alone, the ratio between β major and β minor globin chains was approximately 1.3:1.0, in agreement with previous reports^{3,5}, but addition of 700 $\mu\text{g ml}^{-1}$ AIA to the cultures at the beginning of DMSO induction changed the ratio to 3.4:1.0. AIA at 200 $\mu\text{g ml}^{-1}$ was ineffective in changing the ratio of the β chains and 500 $\mu\text{g ml}^{-1}$ was nearly as effective as 700 $\mu\text{g ml}^{-1}$.

To prove that the identification of globin chains on the basis of their electrophoretic behaviour in acrylamide gels containing Triton X-100 was correct, cytoplasmic lysates of Friend cells were treated with mouse globin antiserum and the resulting immunoprecipitates were analysed by electrophoresis. Of the large numbers of polypeptides shown in Fig. 1, only those bands tentatively identified as globin chains were detectable on acrylamide gels containing a 0–2% (w/v) gradient of Triton X-100. For a more direct comparison of the relative amounts of the two β chains the immunoprecipitates were also electrophoresed in parallel on slab gels containing 2% (w/v) Triton X-100 (Fig. 2). Both β major and β minor globin chains were present in immunoprecipitates of DMSO-treated cells, but in immunoprecipitates of haemin-treated cells only β minor chain

was present. In precipitates of cells induced with DMSO and treated with 700 $\mu\text{g ml}^{-1}$ of AIA at the same time, β major predominated. The relative amount of globin chains in each immunoprecipitate was quantitated from the scans of the X-ray films and found to be in good agreement with the amounts calculated from the fluorographs of the total lysates.

The results reported here indicate that AIA inhibits the increase in intracellular haem in erythroleukaemia cells induced to differentiate by DMSO and reduces the amount of β minor globin chain present. The effect of AIA on erythroleukaemia cells is the opposite of the effect of haemin, which causes an increased amount of β minor chain. As shown in Fig. 3, the unbalanced amount of β -globin chain present after treatment of Friend cells with haemin is reflected also in the types of haemoglobins that are assembled.

Although it has been possible using haemin or AIA to uncouple the expression of the two β -globin genes it has not been possible to dissociate the expression of α -globin from the β -globin genes. In different conditions the ratio between α - and β -globin chains has always been found to be close to one. Only in haemin-treated cells β minor globin chain was somewhat in excess, the α to β ratio being 0.76.

These findings suggest that intracellular haem levels control the expression of only one (β minor) of the two β globin genes and that the expression of the α gene is coupled to the expression of either β gene and not directly dependent on the intracellular level of haem.

Whether expression of the β minor gene is controlled by haem at the level of transcription during the processing of mRNA precursors, at the level of translation, or post-translation remains to be determined. It is known that haemin controls translation of globin chains (see ref. 21 for review). However, pretranslation regulation is also likely, as Ross and Sautner⁷ have reported that globin mRNA increases soon after induction of haemoglobin synthesis by haemin.

We thank Dr W. E. Scott (Hoffman-LaRoche) for the gift of allylisopropylacetamide, and Ms Janet Horn and Christine Marsh for technical assistance. G.R. is a Scholar of the Leukemia Society of America. This research was supported by grants from the US NCI (CA 09171, CA 10815, and CA 19454).

GIOVANNI ROVERA

DAVID ADEN

STUART SURREY

The Wistar Institute of Anatomy and Biology,
36th Street at Spruce,
Philadelphia, Pennsylvania 19104

Received 14 November 1977; accepted 9 January 1978.

- Russell, E. S. & Bernstein, S. E. in *Biology of the Laboratory Mouse* (ed. Green, E. L.) 351 (McGraw-Hill, New York, 1966).
- Gilman, J. G. *Science* **178**, 873–874 (1972).
- Kabat, D., Sherton, C. C., Evans, L. H., Bigley, R. & Koler, R. D. *Cell* **5**, 331–338 (1975).
- Nudel, U. et al. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1100–1104 (1977).
- Rovera, G., Abramczuk, J. & Surrey, S. *FEBS Lett.* **81**, 366–370 (1977).
- Alter, B. P. & Goff, S. C. *Blood* **50**, 867–876 (1977).
- Ross, T. & Sautner, D. *Cell* **8**, 513–520 (1976).
- Marver, H. S., Collins, A., Tschudy, D. P. & Rechigl, M., Jr *J. biol. Chem.* **241**, 4323–4329 (1966).
- Ebert, P. S. & Ikawa, Y. *Proc. Soc. exp. Biol. Med.* **146**, 601–604 (1974).
- Padmanabhan, G., Satyanarayana Rao, M. R. S. & Malathy, K. *Biochem. J.* **134**, 847–857 (1973).
- Friend, C., Scher, W., Holland, J. & Sato, T. *Proc. natn. Acad. Sci. U.S.A.* **68**, 378–382 (1971).
- Rovera, G. & Bonaiuto, J. *Cancer Res.* **36**, 4057–4061 (1976).
- Rovera, G. & Surrey, S. *Cancer Res.* **37**, 4211–4219 (1977).
- Orkin, S. H., Harosi, F. I. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **68**, 378–382 (1975).
- Drysdale, J. W., Righetti, P. & Bunn, H. F. *Biochim. biophys. Acta* **229**, 42–50 (1971).
- Rovera, G., Magarian, C. & Borun, T. W. *Analyt. Biochem.* (in the press).
- Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83–88 (1974).
- Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* **56**, 335–341 (1975).
- Cullen, S. E. & Schwartz, B. D. *J. Immun.* **117**, 136–142 (1976).
- Sassa, S. *J. exp. Med.* **143**, 305–315 (1976).
- Lodish, H. F. A. *Rev. Biochem.* **45**, 39–72 (1976).

Intermediate filaments anchor the nuclei in nuclear monolayers of cultured human fibroblasts

EUKARYOTIC cells display polarity both *in vivo* and *in vitro*, the nucleus being located in a specific region of the cytoplasm. Ultrastructural studies have suggested that various cytoplasmic filamentous structures may be responsible for the anchorage of nuclei in cells *in vivo*^{1–4}. Microfilaments and microtubules are well characterised as cytoskeletal elements of eukaryotic cells^{5,6}. The function of other cytoplasmic filamentous structures, including the intermediate filaments (10 nm filaments)^{7,8} and the ground cytoplasmic filamentous network^{9–11} has not been established. It is also still unclear which of the filamentous structures are responsible for the skeletal function and whether other cytoplasmic structures are involved in the maintenance of the cytoplasmic organisation. Based on ultrastructural, immunofluorescence and biochemical studies, we show here that intermediate filaments mediate the attachment of the nuclei to the growth substratum in the nuclear monolayers of cultured human embryonic fibroblasts. We also suggest that the intermediate filaments have a major role in nuclear anchorage in living cells.

Under phase contrast microscopy, cultured normal human embryonic skin fibroblasts (HES cells) had a distinct nucleus and a granular cytoplasm in which stress fibres⁷ could be seen (Fig. 1a). Cells cultured in the presence of cytochalasin B to disrupt actin-containing microfilaments (CB-HES cells) were typically arborised and showed a bulging of the nucleus^{15,16}, but apparently remained firmly attached to the growth substratum.

To obtain 'nuclear monolayers'^{12,13}, both the HES cells and CB-HES cells were exposed to 0.5% Triton X-100 to disrupt cellular membranes. With CB-HES cells, cytochalasin B was also present in Triton X-100 solution to prevent reorganisation of the actin. Thereafter, low and high ionic actomyosin extraction buffers were used to further disrupt microfilaments and microtubules¹⁴ (see Fig. 1 legend). After this treatment, most of the cytoplasmic structures were solubilised in both types of cells, without loosening the anchorage of the nuclei to the growth substratum, as judged by phase contrast microscopy (Fig. 1b and c). The nuclei could not be removed even by vigorous agitation.

Under indirect immunofluorescence (IFL) microscopy, the distinct filamentous staining which had been noted in untreated HES cells both with human anti-actin antibodies¹⁷ and experimental rabbit anti-actin antibodies¹⁸ was abolished after the treatments described above (Fig. 1e). However, the dense fibrillar fluorescence obtained in untreated HES cells with antibodies specific for intermediate filaments¹⁷ (Fig. 1f) was also seen in extracted nuclear monolayers (Fig. 1g). In contrast to microfilaments, intermediate filaments could still be seen in the CB-HES cells after this treatment (Fig. 1h).

Under electron microscopy, the cells lost their cytoplasmic organelles in both types of cultures after the extraction procedures. Arrays of intermediate filaments, 8–10 nm in diameter, but no microfilaments or microtubules, were seen to traverse the cytoplasmic domain (Fig. 2). Also, the nuclei resisted the treatment and remained as a matrix-like structure lacking nuclear membranes. In tangential sections, the filaments appeared to attach to the nuclear residue, where nuclear pore complexes were also observed (Fig. 3). The extracellular material resisting the extractions was especially distinct after staining with tannic acid (see Fig. 2).

The protein composition of intact cells and of nuclear monolayers of HES cells after actomyosin extractions were compared by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (Fig. 4). In contrast to the untreated cells, the extracted nuclear monolayers gave only a few bands: a

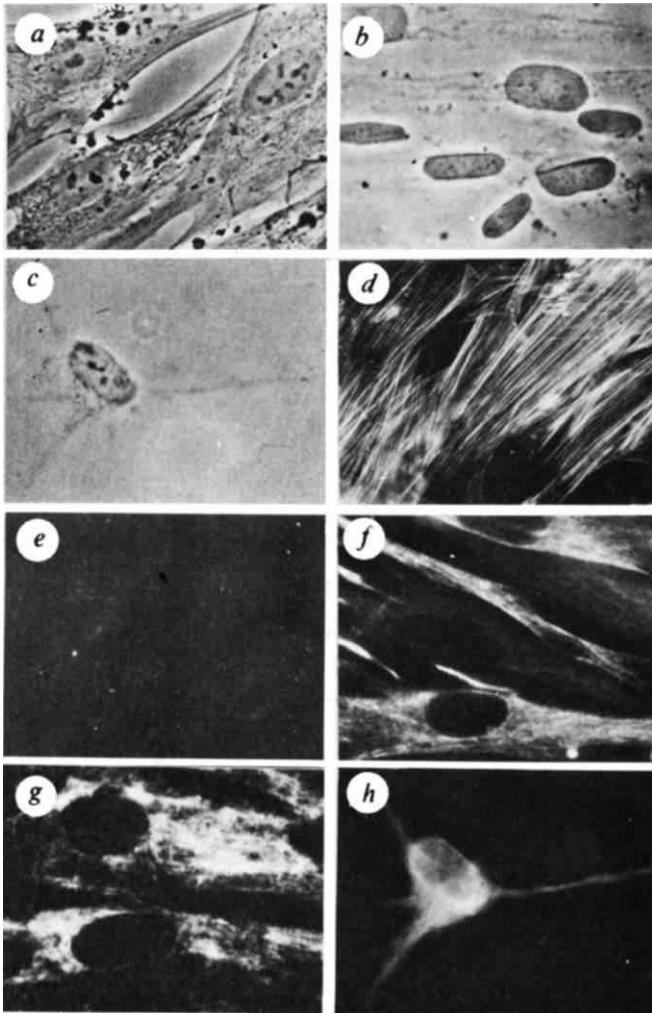


Fig. 1 Human embryonic skin (HES) fibroblasts examined by light microscopy. The cells were cultured in RPMI medium (Flow) supplemented with 10% foetal calf serum (Flow) and antibiotics (HES cells), or in the presence of cytochalasin B (Serva) 10 μg per ml medium for 5 h (CB-HES cells). The cells were cultured for two days until subconfluent. They were then treated with 0.5% Triton X-100 in 40 mM Tris-HCl (pH 7.2) and 1 mM phenylmethylsulphonyl fluoride (Sigma) (for CB-HES cells, the Triton X-100 solution was supplemented with cytochalasin B, 10 μg ml⁻¹), and thereafter according to the procedure of Small and Sobieszek¹⁴ with a low ionic extraction buffer (60 mM KCl, 1 mM EDTA, 2 mM EGTA, 1 mM cysteine, 10 mM ATP and 40 mM imidazole, pH 7), a high ionic extraction buffer (600 mM KCl, 1 mM EDTA, 10 mM ATP, 1 mM cysteine and 40 mM imidazole, pH 7), and again with the low ionic buffer at 22 °C. for 30 min each. For immunofluorescence studies, the cells cultured on glass cover slips were fixed in -20 °C acetone and washed three times in phosphate buffered saline, pH 7.2 (PBS) before staining. Intermediate filaments and actin-containing microfilaments were visualised by indirect immunofluorescence (IFL) techniques as described elsewhere¹⁷. A Zeiss Universal microscope was used for fluorescence microscopy and phase contrast microscopy. *a*, Untreated HES cells showing a dense, granular cytoplasm, a distinct nucleus and stress fibres; *b* and *c*, HES and CB-HES cells, respectively, after treatment with Triton X-100 and actomyosin extraction buffers. Under phase-contrast microscopy, the cells can be seen to have lost all their cytoplasmic structures, but the nuclei are still clearly identifiable, and remain tightly bound to the growth substratum in both figs; *d*, untreated HES cells under IFL microscopy after staining with anti-actin antibodies, and showing a typical filamentous fluorescence; *e*, the cells under IFL microscopy after Triton X-100 and actomyosin extraction treatment, showing loss of fluorescence; *f*, untreated HES cells after treatment with antibodies against intermediate filaments, which are stained as a fine fibrillar network; *g*, HES cells after Triton X-100 and actomyosin extraction treatment still show this staining pattern; *h*, arborised CB-HES cells showing a bright fibrillar fluorescence. Magnification in (a)-(h), $\times 500$.

faint band of molecular weight (MW) 42,000-43,000, representing residual actin¹⁹; a band of 52,000-54,000 corresponding to skeletin or desmin^{14,19}, a subunit protein of intermediate filaments; a band of MW 68,000, probably the major protein of the nuclear protein matrix^{20,21} and a few bands of MW 200,000 which probably include the fibronectin or LETS protein, known to resist the treatments used in this study^{22,23}

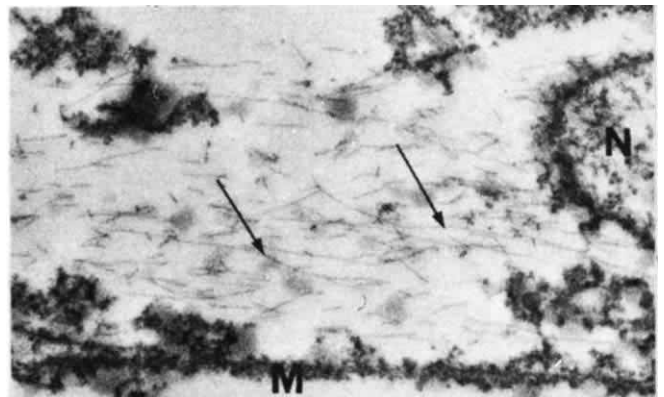
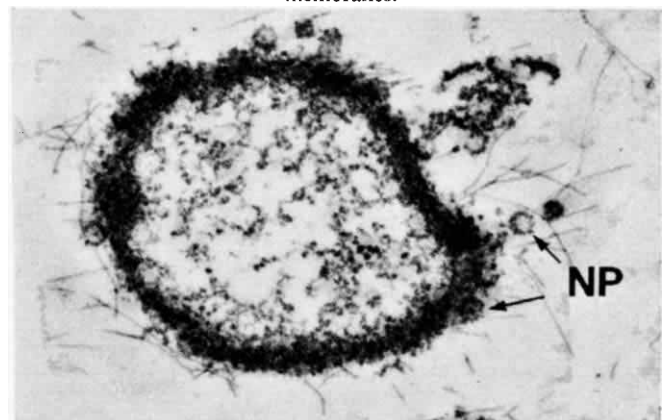


Fig. 2 For electron microscopy studies the cells were treated as in Fig. 1*b*, and thereafter fixed in 2.5% glutaraldehyde, buffered with 0.1 M Na-cacodylate (pH 7.2) for 60 min, and treated with 4% tannic acid in 0.1 M Na-cacodylate for 4 h. The specimens were postfixed in 1% osmium tetroxide for 60 min, dehydrated and embedded in Epon 812. Thin sections were examined with a Jeol 100 B electron microscope at the Department of Electron Microscopy, University of Helsinki. The cells can be seen to have lost almost all the cytoplasmic organelles, only intermediate filaments (arrows), 8-10 nm in diameter, and the nuclear residue (N) being seen. The filaments appear to stretch throughout the cytoplasmic area from the nuclear residue. Extracellular material is seen as an especially dense layer after tannic acid treatment (M). Note the lack of microfilaments and microtubules. Magnification $\times 30,000$.

These findings support the morphological observations that the intermediate filaments were the major cytoskeletal structures present in the extracted nuclear monolayers.

Microtubules have usually been connected with cellular secretory functions²⁴. At least in fibroblasts they have a cytoskeletal function, for instance, in keeping the Golgi apparatus in a juxtanuclear position²⁵. The suggestion of anchorage of the nucleus by microtubules, however, has now been discarded³. In the present study, microtubules could be disrupted without affecting the anchorage of the nuclei in nuclear monolayers. Some studies^{15,16} have suggested a role for microfilaments in nuclear anchorage by showing that cells can be enucleated by centrifugal force after treatment with cytochalasin B. However, this could be due to decreased viscosity of the cortical microfilament layer rather than to disruption of the structures anchoring the nucleus^{15,16}. We could not

Fig. 3 At high magnification ($\times 35,000$) the intermediate filaments can be seen to be attached to the nuclear residue. Nuclear pore complexes are also visible (NP). Note the lack of nuclear membranes.



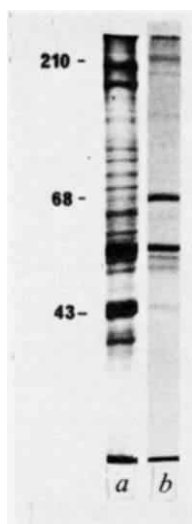


Fig. 4 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of HES cells before (a) and after (b) extraction with detergent and actomyosin extraction buffers. The specimens were solubilised in a sample buffer applied on 7.5% gels. The electrophoresis procedure was carried out according to the method of Laemmli²⁸. Untreated cells showed about 20 bands when stained with Coomassie blue (a). Note the distinct bands of MW ~210,000, 200,000, 52,000–53,000 and 42,000. In extracted nuclear monolayers (b), 2 bands (MW 68,000 and 52,000) are enriched. The residual actin band (MW 42,000) is hardly detectable. Molecular weights ($\times 10^3$) of the marker proteins; thyroglobulin, 210,000; albumin, 68,000; ovalbumin, 43,000, are indicated on the left hand side.

show that microfilaments have a direct role in the nuclear anchorage in nuclear monolayers; disruption of the microfilaments by cytochalasin B treatment and actomyosin extractions did not detach the nuclei. The actin resisting the actomyosin extraction (see Fig. 4), also noted by others^{14,19,27,28}, may represent residues of microfilamentous actin or, more probably, may be an integral part of intermediate filaments, as has been suggested recently^{28,29}.

Apart from intermediate filaments and microfilaments, there are additional cytoplasmic structures like endoplasmic reticulum³⁰, which can contribute to the anchorage of the nucleus in living cells. On the basis of our results, however, indicating that intermediate filaments anchor nuclei in nuclear monolayers, we suggest that they play a major part in nuclear anchorage in intact fibroblasts also. This is in agreement with previous studies assigning a cytoskeletal function to intermediate filaments^{14,28}.

We thank Hannele Laaksonen, Pirkko Leikas-Lazanyi and Raimo Laakia for technical assistance, and Dr Svante Stenman for the cells used in this study, which was supported by the Finnish Medical Research Council, the J. K. Paasikivi Foundation for Cancer Research and the Finnish Culture Foundation.

VELI-PEKKA LEHTO
ISMO VIRTANEN

Department of Pathology, University of Helsinki

PEKKA KURKI

Department of Bacteriology and Immunology,
University of Helsinki, Haartmaninkatu 3, SF-00290,
Helsinki 29, Finland

Received 23 September 1977; accepted 13 January 1978.

- Metuzals, J. & Mushynski, W. E. *J. cell. Biol.* **61**, 701–722 (1974).
- Gray, E. B. & Westrum, L. E. *Cell Tissue Res.* **168**, 445–453 (1976).
- Osborn, M. & Weber, K. *Expl Cell Res.* **106**, 339–349 (1977).
- Bornens, M. *Nature* **270**, 80–82 (1977).
- Pojlard, T. D., Fujiwara, K., Niedermann, R. & Maupin-Szamier, P. in *Cell Motility* (eds Goldman, R., Pollard, T. & Rosenbaum, J.) 689–724 (Cold Spring Harbor Laboratory, 1976).
- Weber, K. in *Cell Motility* (eds Goldman, R., Pollard, T. & Rosenbaum, J.) 403–417 (Cold Spring Harbor Laboratory, 1976).
- Goldman, R. & Knipe, D. M. *Cold Spring Harb. Symp. quant. Biol.* **37**, 523–534 (1973).
- Shelanski, M. L., Yen, S.-H. & Lee, V. M. in *Cell Motility* (eds Goldman, R., Pollard, T. & Rosenbaum, J.) 1007–1020 (Cold Spring Harbor Laboratory, 1976).

- Gray, E. G. *J. Neurocytol.* **4**, 315–339 (1975).
- Wolosewick, J. J. & Porter, K. R. *Am. J. Anat.* **147**, 303–323 (1976).
- Buckley, I. K. *Tissue Cell* **6**, 1–20 (1974).
- Tsai, R. L. & Green, H. *Nature* **243**, 168–170 (1973).
- Bell, D. *Nature* **248**, 505–508 (1974).
- Small, J. W. & Sobieszek, A. *J. Cell Sci.* **23**, 243–268 (1977).
- Shay, J. W., Gershenbaum, M. R. & Porter, K. R. *Expl Cell Res.* **94**, 47–55 (1975).
- Shay, J. W., Porter, K. R. & Prescott, D. M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3059–3063 (1974).
- Kurki, P., Linder, E., Virtanen, I. & Stenman, S. *Nature* **268**, 240–241 (1977).
- Miettinen, A., Virtanen, I. & Linder, E. *J. Cell Sci.* (in the press).
- Lazarides, E. & Hubbard, B. D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4344–4348 (1976).
- Berezney, R. & Coffey, D. S. *Biochem. biophys. Res. Commun.* **60**, 1410–1417 (1974).
- Virtanen, I. *Biochem. biophys. Res. Commun.* **78**, 1411–1417 (1977).
- Brown, S., Levinson, W. & Spudich, J. A. *J. supramolec. Struct.* **5**, 119–130 (1976).
- Hynes, R. O., Destree, A. T. & Mautner, V. J. *J. supramolec. Struct.* **1**, 189–202 (1976).
- Le Marchand, Y., Singh, A., Pazelt, C., Orci, L. & Jeanrenaud, B. in *Microtubules and Microtubule Inhibitors* (eds Burgers, M. & de Brabander, M.) 153–164 (Elsevier, Amsterdam, 1975).
- Moskalewski, S., Thyberg, J., Hinek, A. & Friberg, U. *Tissue Cell* **9**, 185–196 (1977).
- Weber, K. in *Contractile Systems in Non-Muscle Tissues* (eds Perry, S. V., Margreth, A. & Adelstein, R. S.) 23–28 (North Holland, Amsterdam, 1976).
- Izant, J. G. & Lazarides, E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1450–1452 (1977).
- Lazarides, E. & Balzer, D. R., Jr. *J. Cell Biol.* **75**, 255a (1977).
- Buckley, I. K., Raju, T. R. & Stewart, M. J. *J. Cell Biol.* **75**, 257a (1977).
- Franke, W. W. & Kartenbeck, J. in *Progress in Differentiation Research* (ed. Muller-Berat, N.) 213–243 (North Holland, Amsterdam, 1976).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).

Assessment of specificity of oestrogen receptor-DNA interaction by a competitive assay

ALTHOUGH hormonal regulation of gene expression and the nature of specific interactions between steroid hormone receptor and genome have attracted growing interest^{1–5}, very little is known about the molecular nature of receptor-DNA interaction^{4–6}. As a first step in unravelling the specificity and the dynamics of the interaction between uterine oestrogen receptor and DNA^{7–11}, we have selected a well defined model system: a synthetic DNA and DNA-cellulose competition assay. We have found that the AT-rich segment of the double stranded DNA in its intact conformation is required for optimum receptor binding. We have examined oestrogen receptor binding to synthetic DNA with well defined sequences, and determined whether the receptor favours double-stranded or single-stranded regions of the DNA (unwinding protein generally binds preferentially to single-stranded DNA¹²). Finally, we have studied the effect of an intercalating drug, actinomycin D, on the receptor binding.

To study receptor-DNA interaction, a quick, sensitive and versatile binding assay was needed. ³H-oestradiol-receptor complex (³H-OE₂R) binds to DNA-cellulose^{7–11}. We expect, in a competition assay, that soluble DNA will also combine with ³H-receptor complex in such a way that it will compete for the receptor binding to DNA-cellulose pellet. A similar competition assay to study the interaction between DNA and lac repressor^{13,14}, and RNA polymerase¹⁵ using membrane filter and unlabelled DNA compared with ³²P-labelled DNA, has been used quite extensively. While this manuscript was in preparation, a similar DNA-cellulose competition assay had been published for the glucocorticoid receptor¹⁶.

The binding of rabbit uterine oestradiol receptor to DNA cellulose was examined. When increasing concentrations of ³H-OE₂R were incubated at 25 °C for 40 min with fixed amounts of DNA cellulose, an increasing amount of ³H-steroid was bound to DNA cellulose (Fig. 1.) (The problem of saturation¹⁷ is not discussed here.) Only OE₂R complex, but not the free OE₂, bound to DNA-cellulose; in the presence of buffer or BSA (not shown) or with heat-denatured cytosol (Fig. 1), no substantial binding was observed. When DNA-cellulose was replaced by cellulose, very little binding occurred. The difference between radioactivity associated with DNA-cellulose and cellulose was attributed to DNA binding. These experiments suggest that oestradiol combines with the rabbit cytoplasmic receptor(s) in close analogy with the rat and calf system^{2,7,8} and that the hormone-receptor complex then binds to DNA.

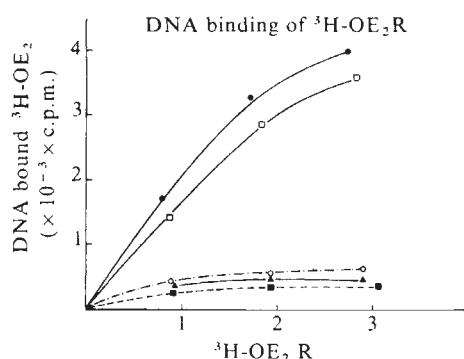


Fig. 1 Binding of OE_2R to DNA-cellulose. Rabbit uterine cytosol (3 mg protein per ml, 105,000g supernatant fraction in Buffer A, containing 50 mM KCl, 20 mM Tris-HCl, pH 7.5, 1 mM EDTA and 1 mM mercaptoethanol, was incubated with 10 nM 3H -oestradiol (96 Ci mmol $^{-1}$) for 60 min at 4°C; the protein-bound oestradiol (3H - OE_2R) was precipitated by ammonium sulphate to 30% saturation. The pellet was dissolved in Buffer A and filtered through Sephadex G-25 (1.25 × 11 cm). DNA-cellulose suspension (200 μ l, containing 7 μ g native calf thymus DNA coupled to cellulose²⁹ in Buffer A) was incubated with increasing amounts of 3H - OE_2R complex (approximately 3,000–9,000 c.p.m. in a total volume of 200 μ l) at 25°C for 45 min with gentle shaking. At the end of the incubation period, the mixture was rapidly chilled, diluted with 0.75 ml cold Buffer A, and centrifuged (1,500g for 7 min). The DNA-cellulose pellet was washed three times with 0.8 ml Buffer A and centrifuged as above; finally, the pellet was assayed for radioactivity by liquid scintillation counting. DNA binding was estimated by subtracting the c.p.m. bound to plain cellulose from the c.p.m. bound to DNA-cellulose. Heat denaturation: the complex was heated for 20 min at 58°C. Each value represents the average of three determinations. ●, OE_2R + DNA-cellulose; □, DNA-bound OE_2R (calculated); ○, OE_2R (heat-inactivated) + DNA-cellulose; ▲, oestradiol alone + DNA-cellulose; ■, OE_2R + cellulose.

When an increasing amount of soluble DNA (calf thymus) was added to a fixed amount of DNA-cellulose and mixed with 3H - OE_2R , the soluble DNA competed and thus decreased the radioactivity in 3H - OE_2R DNA-cellulose pellet (calf in Fig. 2a). The macromolecular structure of the DNA seemed to be required for the competition, as no significant competition was observed with DNase-treated DNA (calf + DNase in Fig. 2a).

We have examined the ability of various natural DNAs to compete with DNA-cellulose for 3H -receptor binding. It can be seen in Fig. 2a that rabbit uterine receptor binds with similar affinity to DNA from rat liver, calf thymus and *Escherichia coli* (salmon sperm DNA was a slightly less effective competitor). tRNA, on the other hand, even at very high concentrations, showed no substantial binding affinity to the receptor. These results suggest that, in general agreement with previous reports^{7–11}, the receptor clearly discriminates between DNA and RNA, but is not specific with respect to DNA base sequences.

Because of the difficulties in demonstrating specificity with natural DNA, we turned to synthetic DNA. If the receptor discriminates among DNA sequences, then we expected that synthetic DNA with different sequences should compete with different efficiency. Consequently, we examined receptor binding to the following synthetic DNAs: poly d(A–T) and poly d(C–G), which have alternating DNA copolymers; and poly d(A) poly d(T) and poly d(A)-poly d(T), which are homopolymers.

Figure 2b shows that the synthetic DNAs had drastically different binding ability for the receptor. Poly d(A–T) was a very effective competitor; poly d(C–G) was much less effective; poly d(A) and poly d(T), single-stranded homopolymers, even at high concentrations, were very poor competitors; and poly d(A)-d(T) was also ineffective (not shown). The receptor seems to bind preferentially to the AT-rich segment of the DNA double helix. A similar preference for AT segments has been reported for other DNA-binding proteins (that is, *E. coli* lac repressor^{14,18,19}, RNA polymerase¹⁵ and polylysine²⁰). We are

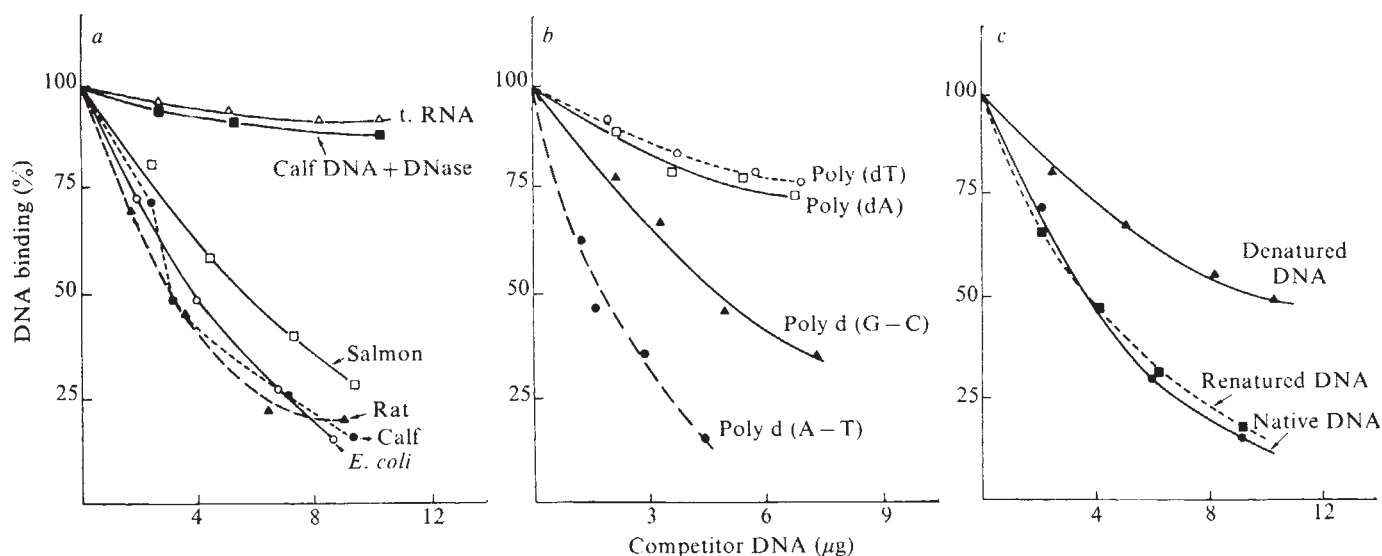


Fig. 2 a, Competition for 3H - OE_2R by soluble DNA. Increasing amounts of soluble calf thymus DNA or other nucleic acids from natural sources (to give the indicated concentration) were mixed with DNA-cellulose (7 μ g DNA) in 200 μ l Buffer A, and 200 μ l of partially purified 3H - OE_2R was added (as described in Fig. 1). The mixture was incubated for 2 h at 25°C with gentle shaking. Samples were then chilled, centrifuged, and the DNA-bound radioactivity in the pellet determined as described in Fig. 1. Results are expressed as a percentage of binding without soluble DNA (100% value was about 3,400–3,600 c.p.m. per 7 μ g DNA; triplicate determinations). DNase digestion: Soluble DNA was incubated at 30°C for 35 min with pancreatic DNase (Worthington, final concentration 50 μ g ml $^{-1}$ in 50 mM Tris-HCl, pH 7.5, containing 1.5 mM CaCl $_2$ and 1.5 mM MgCl $_2$) and the reaction was terminated by addition of EDTA to a final concentration of 5 mM. For the DNase control, EDTA was added before the DNase. DNA was sheared by 10 passages through a 27-gauge needle. **b**, Competition for 3H - OE_2R by soluble synthetic polynucleotides (poly d(A–T), poly d(C–G), poly d(A) and poly d(T) from PL-Biochemicals). DNA-cellulose (0.2 ml; 7 μ g DNA) containing increasing amounts of soluble synthetic polynucleotide, was incubated with 0.2 ml 3H - OE_2R for 2 h at 25°C as described above. DNA-bound radioactivity was assayed as above, and expressed as percentage of binding without competitor (100% value was about 3,500–3,600 c.p.m. per 7 μ g DNA; triplicate determinations). **c**, Competition for 3H - OE_2R by native and denatured DNA. Increasing amounts of native or denatured calf thymus DNA competitor were mixed with DNA-cellulose and then incubated in a final 0.4 ml reaction volume, with 3H - OE_2R as described above. The DNA binding was assayed and expressed as in **a**. The calf thymus DNA (200 μ g ml $^{-1}$ Buffer A) was denatured by heating for 5–10 min at 100°C and rapidly cooled in ice (ΔA_{260} = 32%). Heat denatured DNA was renatured by slow cooling at 55°C for 18 h.

testing several natural DNAs in an attempt to verify the correlation between AT content and oestrogen receptor binding. In contrast to oestrogen receptor, binding studies with glucocorticoid^{16,21} and thyroid hormone²² receptors have not revealed any preferential binding sites on the DNA; also, glucocorticoid receptor complex binds less well to synthetic than to natural DNA¹⁶.

We have, however, considered the possibility that soluble DNA might cause receptor aggregation⁸ and retention by DNA-cellulose. Although aggregation could not be ruled out, this seems unlikely, as an increasing amount of soluble DNA effectively inhibits the pelleting of the ³H-OE₂R with DNA-cellulose. The addition of soluble DNA just before centrifugation did not affect the binding to DNA cellulose; soluble DNA did not interfere with the separation of the bound form. Finally, the possibility that some of the differences observed in natural DNA binding may be due to differences in the size of DNA molecules seems unlikely, as DNA sonication had very little effect on binding (data not shown).

If the receptor is an unwinding protein, it would bind preferentially to single-stranded denatured DNA¹². We have compared receptor binding to double-stranded with that to denatured DNA. Competition experiments (Fig. 2c) demonstrated that heat denaturation (100 °C for 5 min) of calf thymus

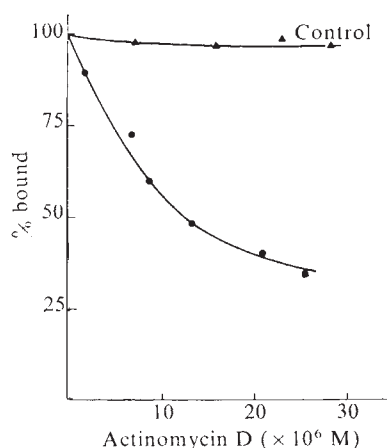


Fig. 3 Inhibition of receptor-DNA binding by actinomycin D. DNA-cellulose (0.2 ml; 7 µg DNA) in the presence of increasing concentrations of actinomycin D, was incubated with 0.2 ml ³H-OE₂R complex for 45 min at 25 °C. DNA-bound radioactivity was assayed as described above. Results are expressed as a percentage of binding without drug (100% value was about 3,500–3,600 c.p.m. per 7 µg DNA; triplicate determinations).

DNA substantially reduced its binding capacity, suggesting that double-stranded native DNA is required for the optimum binding. To substantiate this conclusion, heat-denatured DNA was renatured with complete recovery of the competing activity. The receptor is therefore probably not an unwinding protein. Our findings are in general agreement with those of Rochefort (ref. 23 and H. Rochefort, personal communication). A similar conclusion was reached for the *lac* repressor^{14,18,19}.

Actinomycin D is known to intercalate into the minor groove of DNA with a preference for guanine²⁴. Figure 3 shows that actinomycin D inhibits the binding of oestrogen receptor to DNA-cellulose; 50% inhibition was observed with 1.2×10^{-5} M actinomycin. The possibility that the drug somehow modifies the oestrogen receptor seems unlikely. No inhibition of binding was observed when the receptor complex was incubated with the drug only, separated from the actinomycin D by G-25 filtration and then incubated with DNA-cellulose (labelled control in Fig. 3). This experiment suggests that actinomycin D inhibits receptor binding to DNA rather than affecting receptor stability. Ethidium bromide, another intercalating drug, also inhibits the oestradiol receptor-DNA interaction²³. These intercalation studies suggest that the base pairs in the DNA helix

(and/or their conformation) might be directly involved in the receptor binding.

Oestrogen receptor interacts nonspecifically with almost any DNA^{7–11}. However, the present DNA-cellulose competition study revealed that (1) the receptor seems to bind preferentially to AT-rich segments of the double-stranded DNA; and (2) the receptor is probably not an unwinding protein. Moreover, the intact conformation of the DNA double helix is probably required for optimum receptor binding, since it is inhibited by intercalating drugs, such as actinomycin D and ethidium bromide²³.

It is not yet clear how chromosomal proteins²⁵ and the organisation of chromosomal subunits affect receptor-genome recognition. Our findings, and those of Andre *et al.*²³ with ethidium bromide, are consistent with the view that, besides the nonspecific electrostatic interactions with phosphate groups in DNA, the receptor can recognise and form a contact with the nucleotide bases through the major or minor grooves of the DNA helix^{14,20,26}. We have studied the DNA recognition by substituting the bromine for the 5-methyl group of thymine in the major groove by incorporating bromouracil into DNA^{27,28}, and found that the receptor binds tenfold more tightly to bromouracil-substituted DNA than to unsubstituted DNA (J. K. and T. Fazy, unpublished results). The particular value of the competitive DNA-cellulose binding assay is to provide a simple chemical probe for receptor site mapping. The present study can only be considered preliminary until purified receptor is available.

This investigation was supported by NCI grants CA 10064, P 30 14194 and RR 5589. We thank R. Axel, C. Cantor, T. Fazy, C. Monder, P. van Hippel and K. Yamamoto for their helpful criticism and stimulating discussion.

JOHN KALLOS

VINCENT P. HOLLANDER

Research Institute of the Hospital for Joint Diseases,
and Medical Center, Mount Sinai School of Medicine,
1919 Madison Avenue, New York, New York 10035

Received 3 November; accepted 28 December 1977.

- King, R. J. B. & Mainwaring, W. I. P. *Steroid-Cell Interactions* (University Park Press, Baltimore, 1974).
- Jensen, E. V. & DeSombre, E. R. *A. Rev. Biochem.* **41**, 203–230 (1972).
- Buller, R. E. & O'Malley, B. W. *Biochem. Pharmacol.* **25**, 1–12 (1976).
- Gorski, J. & Gannon, F. A. *Rev. Physiol.* **38**, 425–450 (1976).
- Yamamoto, K. R. & Alberts, B. M. *A. Rev. Biochem.* **45**, 721–746 (1976).
- Kallos, J. *Nature* **265**, 705–710 (1977).
- Yamamoto, K. R. & Alberts, B. *Cell* **4**, 301–310 (1975).
- Yamamoto, K. R. & Alberts, B. *J. biol. Chem.* **249**, 7076–7086 (1974).
- Andre, J. & Rochefort, M. *FEBS Lett.* **50**, 319 (1975).
- King, R. J. & Gordon, J. *J. Steroid Biochem.* **240**, 185–187 (1972).
- Toft, D. O. *J. Steroid Biochem.* **3**, 515–522 (1972).
- Alberts, B. M. & Frey, L., *Nature* **227**, 1313–1318 (1970).
- Bourgeois, S. & Riggs, A. D., *Biochem. biophys. Res. Commun.* **38**, 348–353 (1970).
- Bourgeois, S. & Pfahl, M. *Adv. Protein Chem.* **30**, 1–93 (1976).
- Chamberlin, M. J. *A. Rev. Biochem.* **43**, 721–775 (1974).
- Simons, S. S., Jr *Biochim. biophys. Acta* **496**, 349–359 (1977).
- Chamness, G. C., Jennings, A. S. & McGuire, W. L. *Biochemistry* **13**, 327–331 (1974).
- Von Hippel, P. H., Revzin, A., Gross, C. A. & Wang, A. C. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4808–4812 (1974).
- Von Hippel, P. H., Revzin, A., Gross, C. A. & Wang, A. C. in *Protein-Ligand Interactions* (de Gruyter, Berlin, 1975).
- Von Hippel, P. H. & McGhee, J. D. *A. Rev. Biochem.* **41**, 231–300 (1972).
- Simons, S. S., Jr, Martinez, H. M., Garcea, R. L., Baxter, J. D. & Tomkins, G. M. *J. biol. Chem.* **251**, 334–343 (1976).
- MacLeod, K. M. & Baxter, J. D., *J. biol. Chem.* **251**, 7380–7387 (1976).
- Andre, J., Pfeiffer, A. & Rochefort, H. *Biochemistry* **15**, 2964–2969 (1976).
- Sobell, H. M. & Jain, S. C. *J. molec. Biol.* **68**, 21–34 (1972).
- Spelsberg, T. C. in *Acidic Proteins of the Nucleus* (ed. Cameron, J. L. & Jeter, J. R., Jr) (Academic, New York, 1974).
- Jovin, T. M. *A. Rev. Biochem.* **45**, 889–920 (1976).
- Lin, S. Y. & Riggs, A. D. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2574–2576 (1972).
- Ogata, R. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4973–4976 (1977).
- Alberts, B. & Herrick, G. *Methods Enzymol.* **21D**, 198–217 (1971).

New cluster of ribosomal genes in *Bacillus subtilis* with regulatory role in sporulation

STUDIES of the onset of sporulation in *Bacillus subtilis* can provide information of value to the study of differentiation in general. Little is known about the molecular mechanisms involved in initiating sporulation, or any other developmental process. Mutations at any one of nine distinct genetic loci may

block sporulation in *B. subtilis* at early stages without affecting vegetative growth^{1,2}. These mutations are pleiotropic, in that they are not only blocked in sporulation but are also impaired, to various degrees, in stage 0 functions such as antibiotic and protease production, competence for transformation, and sensitivity to surface active antibiotics and phage³. We have examined the ribosomal composition of these stage 0 mutants and of the same mutants bearing an additional mutation in another locus, *abrB*, that results in suppression of the pleiotropic mutant phenotype without suppressing the sporulation defect. We report here that *abrB* mutations define a new cluster of ribosomal protein genes, with the origin proximal to the main group of ribosomal genes. These findings demonstrate the involvement of ribosomes in the control of early stages of the sporulation cycle.

The phenotypes of strains bearing one of the most pleiotropic stage 0 mutations, *spo0A*, together with an unlinked suppressor, *abrB*, are shown in Table 1. Of the 20 suppressed *spo0A* derivatives that we have studied, 17 were similar to those described in Table 1. They were all resistant to polymyxin and wild-type antibiotic, they produced extracellular proteases, and some of them produced antibiotic. Additionally, they were non-permissive for two bacteriophages that only replicate in stage 0 mutants (Table 1). However, the ability to sporulate

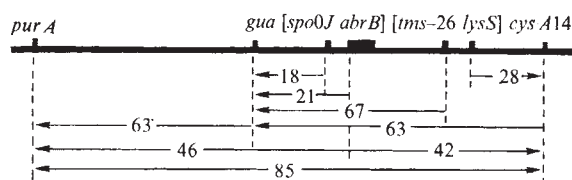


Fig. 1 A genetic map of the replication origin region of the *B. subtilis* genome, constructed from PBS1 transduction crosses⁸. Details of mapping of *gua* and *spo0J* will be published elsewhere (J.T. & J.A.H., in preparation). The order of the markers in brackets has not been determined by three-factor crosses. The figures are of 100-% cotransduction and the head of the arrows point to the selected markers. *abrB* is shown as a single locus and the results for all the *abrB* markers have been averaged for the purposes of this map, although genes for several ribosomal proteins are probably located here. The origin of replication is apparently located half-way between *purA* and *cysA* (ref. 11).

was not recovered. The phenotypes of the *abrB* mutants were similar to some partial revertants of stage 0 sporulation mutants that have been isolated before, but were not mapped^{4,5}. All the *abrB* mutants were located near the origin of replication on the genetic map, to the left of *cysA14* (Fig. 1). The main group of ribosomal genes is to the right of *cysA14* (ref. 6).

The ribosomal subunit composition of stage 0 mutants and their derivatives carrying *abrB* markers was determined by

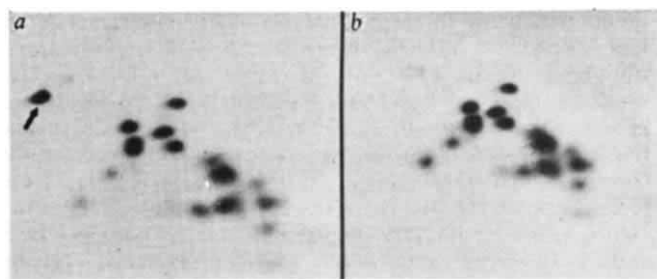


Fig. 2 Comparison of the 50S core proteins of a stage 0 mutant and a partially suppressed derivative. The cells were grown to 2 h post-exponential phase. Ribosomal subunits were isolated by centrifugation in 10–40% sucrose gradients after a 0.5 M KCl wash, according to the methods of Freinstein and Blobel¹². Ribosomal proteins were extracted from the isolated subunits with acetic acid, dialysed against distilled water, and subjected to electrophoresis in two dimensions, using a modified method of Martini and Gould¹³. The first dimension gel was 5% polyacrylamide, electrophoresed at pH 4.5; the second dimension was 10% polyacrylamide with sodium dodecyl sulphate, pH 7.6. a, Strain JH646, *spo0A12*; b, strain 646MS, *spo0A12 abrB15*. The arrow points to a protein, tentatively identified as L2, because of its polarity and high molecular weight¹⁴.

two-dimensional polyacrylamide gel electrophoresis. Strain 646MS was found to be completely devoid of protein L2 of the 50S subunit (Fig. 2). Similarly, strain R15-12 had another low molecular weight 50S protein missing, and strain 703-3 had a charge alteration in a 50S protein. Neither of the altered proteins of these latter strains has been identified. We are now studying the ribosomal protein changes in 14 additional *abrB* mutants. The eight mutants that we have studied so far all have alterations in different ribosomal proteins. These different, specific changes were consistently found in ribosomal proteins isolated from cells either in the log phase, or stationary phase, of growth. Ribosomal core proteins of the parental strain, JH642, and all the other stage 0 mutants did not differ from those of sporulating strains (M.S. & J.A.H., in preparation).

The data show that the *abrB* region probably consists of a group of structural genes specifying several ribosomal proteins. Alteration of these proteins permits the expression of the phenotypes masked in the stage 0 mutants. This result provides a clue to controls that are operating at the molecular level early in sporulation. Presumably, species of mRNA that specify proteins concerned with proteases, antibiotics and other early functions in sporulating cells are not translated in stage 0 mutants. Ribosomal alterations (*abr* mutations) permit these messages to be read, without permitting sporulation to take place. In normal sporulating strains, this mRNA may be read when a change takes place on the ribosome (or possibly, on

Table 1 Phenotypes of some partial revertants of *spo0A* mutants

Strains	Relevant genotype	Polymyxin resistance	Competence	Antibiotic production	Antibiotic resistance	Protease production	Sporulation	Bacteriophage sensitivity EOP			
								Φ2	SP01	Φ15	Φ29
JH642	<i>spo0A</i> ⁺	+	++	+	+	++	+	R	0.1	R	0.3
JH646	<i>spo0A12</i>	—	+	—	—	+	—	1.0	1.0	1.0	1.0
R15-12	<i>spo0A12 abrB22</i>	+	++	—	—	++	—	R	0.5	0.003	0.2
703-3	<i>spo0A677 abrB3</i>	+	++	+	+	++	—	R	0.4	0.002	0.2
646MS	<i>spo0A12 abrB15</i>	+	++	—	+	++	—	R	0.3	R	0.1

Strain R15-12 was selected as a phage Φ15 resistant colony, 703-3 as antibiotic-resistant. 646MS was a spontaneous mutant that was not selected for. Polymyxin resistance; +, growth on 70 μg ml⁻¹ polymyxin in solidified minimal medium; —, no growth. Competence; ++, transformation frequency > 10⁻⁴; +, > 10⁻⁵ Trp⁺ transformants per recipient cell. Antibiotic production was determined on plates with a soft agar overlay containing the indicator strain, *B. amyloliquefaciens* H. Antibiotic resistance was determined in a similar manner, using the tests strains as indicators. Protease production; the notation reflects halo widths on Schaeffer's sporulation medium containing 0.5% casein. Sporulation frequency was determined by plating heat treated (80 °C, 10 min) cultures onto nutrient agar plates. Phage sensitivity; EOP (efficiency of plating) was calculated as the number of phage plaques, normalised to the number of plaques on strain JH646; R, phage resistant.

the mRNA) early in sporulation. In other words, one of the functions of the stage 0 gene products could be to alter the ribosomes so that they may recognise and translate certain species of message. Alternatively, the products are required to process the messages to enable them to be read by normal ribosomes. As mentioned above, we have found no changes in core ribosomal proteins from any of the stage 0 mutants, but recent evidence indicates that changes do take place in ribosome-associated molecules, early in sporulation (refs 7–10 and M.S. and J.A.H., in preparation). Of course, we cannot rule out the possibility that we are observing a novel control process, mediated by ribosomes, and not directly concerned with translation.

We thank Shu Mei Chen and Helen Coukoulis for technical assistance.

JOHN TROWSDALE
MAUREEN SHIFLETT
JAMES A. HOCH

Department of Microbiology,
Scripps Clinic and Research Foundation,
La Jolla, California 92037

Received 12 September 1977; accepted 6 January 1978.

1. Hoch, J. A. *Adv. Genet.* **18**, 69–99 (1976).
2. Pigot, P. J. & Coote, J. G. *Bact. Rev.* **40**, 908–962 (1976).
3. Brehm, S. P., Staal, S. P. & Hoch, J. A. *J. Bact.* **115**, 1063–1070 (1973).
4. Guespin-Michel, J. F. *Molec. gen. Genet.* **112**, 243–254 (1971).
5. Ito, J. *Molec. gen. Genet.* **124**, 97–106 (1973).
6. Young, F. E. & Wilson, G. A. in *Spores* Vol. 6 (eds Gerhardt, P., Costilow, R. N. & Sadoff, H. L.) 596–614 (American Society for Microbiology, Washington, D.C., 1975).
7. Domoto, T., Kobayashi, K. & Kobayashi, Y. in *Spores* Vol. 6 (eds Gerhardt, P., Costilow, R. N. & Sadoff, H. L.) 307–313 (American Society for Microbiology, Washington, D.C., 1975).
8. Fortnagel, P. & Bergman, R. *Biochim. biophys. Acta* **299**, 136–141 (1973).
9. Leighton, T. J. *J. molec. Biol.* **86**, 855–863 (1974).
10. Rhase, H.-J., Hoch, J. A. & Grosscurth, R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1125–1129 (1977).
11. Harford, N. et al. in *Microbiology—1976* (ed. Schlessinger, D.) 28–34 (American Society for Microbiology, Washington, D.C., 1976).
12. Freinstein, C. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3392–3396 (1975).
13. Martini, O. H. W. & Gould, H. J. *J. molec. Biol.* **62**, 403–405 (1971).
14. Geisser, M., Tischendorf, G. W. & Stöffler, G. *Molec. gen. Genet.* **127**, 129–145 (1973).

RNase H hydrolysis of the 5' terminus of the avian sarcoma virus genome during reverse transcription

NUCLEOTIDE sequence analyses of the ends of the avian retrovirus genome^{1–6} have confirmed predictions of the terminally redundant nature of the viral RNA^{7,8}. These studies have also provided the basis for several models of retrovirus proviral DNA synthesis, all of which have implicated the RNase H activity associated with viral reverse transcriptase in a functional role in the continued and uninterrupted DNA transcription of the viral RNA genome^{1–8}. According to these models DNA synthesis initiates on the tRNA^{trp} primer molecule located close to the 5' end of the viral genome^{9,10} and transcription proceeds to the terminus. Presumably, at this time a suitable substrate is available for the retrovirus RNase H, which is a processive exoribonuclease requiring an unblocked terminus of the RNA moiety of RNA:DNA hybrids for activity^{11–13}. Release of the terminally repeated nucleotides from the hybrid region of the viral genome would create 'sticky ends' with the tRNA^{trp}-initiated DNA at the 5' end capable of hybridising to the terminally redundant RNA genomic sequences at the 3' end of the same or a second 35S RNA subunit. This latter reaction would facilitate uninterrupted transcription from the 5' to the 3' end of the viral genome resulting in genome-length DNA transcripts^{15–20}. We have recently obtained evidence indicating that DNA transcripts much longer than the distance between the tRNA^{trp} primer molecule and the 5' end of the viral genome and containing nucleotide sequences representing the 3' region of the viral genome can be synthesised by the reverse transcriptase *in vitro*²¹. Thus, it seems that DNA synthesis initiated at the 5' end of the viral genome continues at the 3' end in enzymatic reactions *in vitro*. If the avian sarcoma virus (ASV) reverse transcriptase-associated RNase H

activity is required for the continued transcription of the viral genome at the 3' end as proposed in above models, then release of 5' terminally-located ribonucleotides should be apparent during DNA synthesis *in vitro*. We report here that ribonucleotides are indeed released from the viral RNA genome during reverse transcription, and that hydrolysis occurs at a specific site near the 5' terminus. These studies exemplify the first demonstration of RNase H hydrolysis occurring during reverse transcription of the retro-virus RNA genome *in vitro* and implicate functional role for this activity during replication of retroviruses.

Previous studies of RNase H activity during reverse transcription *in vitro* failed to detect an effect on either DNA synthesis or the nature of the DNA product^{22,23}. The reason for this became evident when it was shown that the bulk of the DNA synthesised from the retrovirus genome contained nucleotide sequences representing approximately 1% of the viral genomic sequences^{5,19,20}. In all of the analyses performed in the previous studies the assay systems were not rigorous enough to detect an effect of RNase H on a small portion (<10%) of the viral RNA. We have therefore devised experiments sufficiently sensitive to detect RNase hydrolysis of less than 1% of the viral genome even in conditions in which less than 10% of the viral RNA molecules present in the reaction mixture are utilised as template primers. The impetus for these studies was provided by preliminary observations on the nature of DNA products synthesised in reconstructed reactions containing 35S RNA·tRNA^{trp} complexes as template·primers for reverse transcription. Analysis of the tRNA^{trp}-initiated DNA product with the single-strand specific S₁ nuclease indicated that the DNA transcripts became sensitive to the enzyme with time of incubation (data not shown). However, since these results could be explained by either RNase H hydrolysis of the genomic RNA complexed with the DNA transcripts, or a strand-displacement mechanism which somehow renders the DNA transcripts sensitive to a single-strand nuclease, experiments were performed to directly determine whether ribonucleotides were removed from the RNA template during DNA synthesis. In these experiments ³²P-labelled viral RNA was used in reactions containing the purified RNA-directed DNA polymerase. After a brief period of DNA synthesis (30 min) the reaction was terminated with EDTA

Table 1 Recovery of RNase T₁ oligonucleotides*

Oligonucleotide number	c.p.m. recovered in oligonucleotides		Ratio c.p.m.(+dNTP)/c.p.m.(–dNTP)
	–dNTP	+dNTP	
2	427	469	1.10
3	457	516	1.13
4	247	292	1.18
308	480	455	0.95
7	611	587	0.96
10	217	244	1.12
11	440	453	1.03
12	149	188	1.26
13	791	154	0.19
14	525	474	0.90
15	498	442	0.89
17	323	292	0.90
20	659	809	1.23
21	442	573	1.30
22	193	285	1.48
23	317	250	0.79
30	523	409	0.78
101	817	889	1.09
102	452	477	1.06
109	329	338	1.03
110	308	281	0.91
111	351	427	1.22
403	351	438	1.25
			Mean† = 1.07 ± 0.18

*Oligonucleotides resolved in Fig. 2a (–dNTP) and Fig. 2b (+dNTP) were cut out and the radioactivity measured as described previously⁴.

†Mean value for all oligonucleotides except T–13.

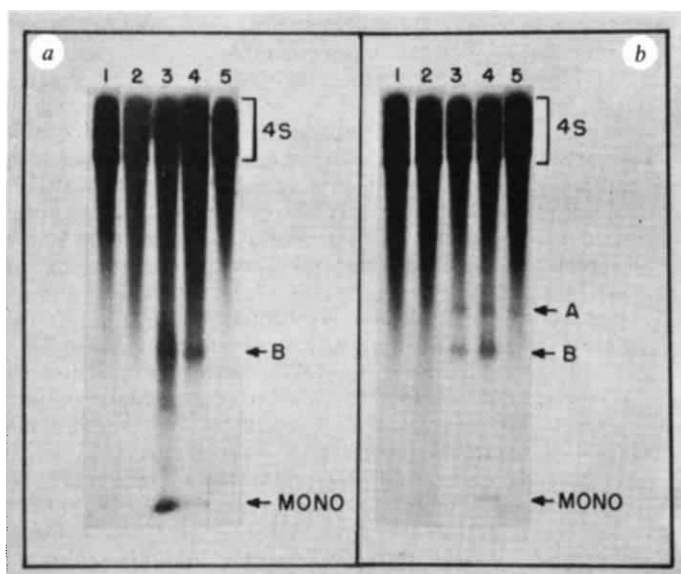


Fig. 1 Polyacrylamide gel electrophoresis of ^{32}P -labelled ASV RNA and ASV RNA:DNA hybrids after incubation with the purified RNA-directed DNA polymerase. Freshly transformed cultures of ASV (strain B77) infected duck cells were washed several times with phosphate-free growth medium, and then labelled with phosphate-free medium containing 1 mCi ml^{-1} carrier-free ^{32}P -orthophosphate (New England Nuclear) and 2% dialysed calf serum. Culture fluids were collected at 12 h intervals, clarified, and stored at 4°C . The virus was pelleted from the pooled culture fluid, treated with Pronase-sodium dodecyl sulphate-phenol, and the 70S RNA isolated by zonal centrifugation. The ^{32}P -labelled 70S RNA had a specific activity of approximately 5×10^6 c.p.m. μg^{-1} . **a**, Release of oligonucleotides from viral RNA during DNA synthesis. Reaction mixtures (0.01 M Tris-HCl, pH 8.1; 0.01 M MgCl_2 ; 1.4% (v/v) 2-mercaptoethanol; 0.03 M NaCl; $250\text{ }\mu\text{g ml}^{-1}$ yeast tRNA) containing $3\text{--}4 \times 10^6$ c.p.m. of ^{32}P -labelled viral RNA⁶ were incubated for 30 min at 37°C alone (1) and in the presence of the following components: (2) 12 U of $\alpha\beta$ AMV DNA polymerase; (3) 60 μM of each of the four deoxynucleoside triphosphates (dNTPs), 12 U of $\alpha\beta$ AMV DNA polymerase; (4) 1 mM of each of the four dNTPs, 12 U of $\alpha\beta$ AMV DNA polymerase; and (5) as in (3) with the addition of $10\text{ }\mu\text{g ml}^{-1}$ poly (dC)-oligo (dG)₁₂₋₁₈. The DNA polymerase was present in saturating amounts for the enzymatic synthesis of DNA. Incubations were terminated by the addition of EDTA to 20 mM and sodium dodecyl sulphate to 1%. After heating to 100°C for 1 min, the samples were subjected to polyacrylamide slab gel electrophoresis at 400 V for 6.5 h at 10°C . Slab gels were composed of 3 cm of 3% polyacrylamide and 27 cm of 20% polyacrylamide in Tris-borate buffer, pH 8.3. The 3% polyacrylamide portion was removed and the remainder of the gel was subjected to autoradiography. In the experiments in which DNA was synthesised 0.7–1.1% of the total c.p.m. of input ^{32}P -labelled viral RNA was released as oligonucleotides. **b**, Release of oligonucleotides from preformed viral RNA:DNA₁₀₀ hybrids after incubation with reverse transcriptase. Two specific DNA transcripts initiated from the tRNA^{trp} primer molecule and representing genomic sequences of the region of the viral RNA between the tRNA^{trp} molecule and the 5' terminus were synthesised from 35S RNA-tRNA^{trp} template-primer complexes in reconstructed enzymatic reactions, isolated by polyacrylamide-formamide gel electrophoresis, and purified as previously reported¹⁹. The initial DNA transcript of 100 nucleotides in length (iDNA₁₀₀) represents the genomic sequences from the 3' OH of tRNA^{trp} directly to the 5' end of the viral RNA^{4,5,14}. iDNA₇₀ represents genomic sequences up to 70 nucleotides distal from tRNA^{trp}; these transcripts do not extend to the 5' terminus of the viral genome^{4,14}. ^{32}P -labelled viral RNA (3×10^6 c.p.m.) was annealed alone (slot 1) or in the presence of 25 ng of iDNA₇₀ (slot 2) or iDNA₁₀₀ (slots 3, 4, 5) in $1\text{ }\mu\text{l}$ of 0.6 M NaCl, 0.02 M Tris-HCl, pH 7.4, 0.01 M EDTA, 400 $\mu\text{g ml}^{-1}$ yeast tRNA for 30 min at 68°C . The annealed mixtures were then incubated in $20\text{ }\mu\text{l}$ of reaction mixture (minus NaCl) with 12 U of $\alpha\beta$ AMV DNA polymerase. iDNA₁₀₀: ^{32}P -labelled RNA hybrid complexes were also incubated in the presence of 60 μM of each of the four unlabelled dNTPs (slot 4), or in the presence of $10\text{ }\mu\text{g ml}^{-1}$ poly(dC)-oligo(dG)₁₂₋₁₈ (slot 5). Mono, mononucleotides.

and sodium dodecyl sulphate; the products were denatured and then applied to a composite polyacrylamide gel capable of resolving 35S RNA from smaller oligonucleotides of 1–75 nucleotides in length. Control reactions containing ^{32}P -labelled viral RNA and reverse transcriptase but lacking deoxynucleoside triphosphates (dNTPs) were also analysed by this procedure. As shown by the autoradiogram presented in Fig. 1a, specific oligonucleotides can be resolved in the gel after DNA synthesis. The appearance of these oligonucleotides was dependent on DNA synthesis, since oligonucleotides were not observed when ^{32}P -labelled viral RNA was incubated in the presence of enzyme but in the absence of DNA synthesis. Increasing the dNTP concentration to levels that facilitate synthesis of longer tRNA^{trp}-initiated DNA transcripts¹⁹ did not significantly increase the amount of oligonucleotides released. Approximately 1% of the total ^{32}P radioactivity was released during DNA synthesis, and 60% of that released was recovered in a specific large oligonucleotide indicated by band B in Fig. 1a. The release of oligonucleotides during DNA synthesis seemed to be a function of the viral RNase H activity since specific inhibitors of this enzyme²⁴ inhibited the release of oligonucleotides during DNA synthesis (Fig. 1a).

In an effort to determine if the 5' terminus of the ASV genome can provide a substrate for RNase H when complexed to DNA transcripts, we have analysed the release of oligonucleotides from defined reconstructed hybrid structures at the 5' end of the viral genome. As seen in Fig. 1b, iDNA₁₀₀, which is the complete 5' terminal containing nucleotide sequences complementary to the entire distance between the tRNA^{trp} primer and the end of the genome, rendered the viral genome sensitive to viral RNase H. One of oligonucleotides released migrated in the same position as band B released during DNA synthesis from tRNA^{trp}. ^{32}P -labelled viral RNA complexes (Fig. 1a). Inhibition of the viral RNase H activity prevented the release of these oligonucleotides from the reconstructed DNA:RNA hybrid structures. If DNA transcripts shorter than the distance between the tRNA^{trp} primer and the 5' end of the viral genome (iDNA₇₀) were reannealed to ^{32}P -labelled viral RNA, and the complex incubated with reverse transcriptase, no release of oligonucleotides could be detected. However, both DNA hybrid structures provide adequate substrates for *Escherichia coli* RNase H which is an endoribonuclease, indicating that suitable RNA:DNA hybrids were formed with both species of DNA transcripts during the reannealing reaction (data not shown).

To demonstrate unequivocally the release of 5' terminal oligonucleotides during DNA synthesis, ^{32}P -labelled viral RNA was incubated with RNA-directed DNA polymerase as described in Fig. 1a legend. After synthesis, the labelled RNA was denatured, digested to completion with RNase T₁

Table 2 Nucleotide composition of bands A and B

Expt 1 Products of RNase U digestion*	
Band A:	U ₄ A, 7 m GpppG ^m CCA, CCA
Band B:	U ₄ A, 7 m GpppG ^m CCA
Oligonucleotide T-13†:	U ₄ A, 7M GpppG ^m CCA, 2CCA, U ₂ CA, CA, U ₂ G
Expt 2. Products of alkaline phosphatase plus RNase T ₁ digestion‡	
Band B:	7mGpppG ^m C _{OH} + Pi
Oligonucleotide T-13:	7mGpppG ^m C _{OH} + Pi
Expt 3. Products of RNase A plus RNase T ₁ digestion§	
Band B:	7mGpppG ^m C, AU, AC, 2C, 2U
Oligonucleotide T-13:	7mGpppG ^m C, 3AU, 3AC, 5C, 5U,G

*RNase U₂ digestion was carried out as described by Cashion *et al.*²⁵

†The sequence of oligonucleotide T₁₃ is: mGpppG^mCCAUUU ACCAUUCACCACAUUG. The terminal redundancy^{1,2} present at the 3' end of PrC RNA is underlined.

‡The presence of cap was confirmed by electrophoresis on 3 MM paper as described by Furuichi *et al.*²²

§Products of RNase A plus RNase T₁ digestion were characterised as described previously⁴.

and the resulting oligonucleotides were resolved by two-dimensional polyacrylamide gel electrophoresis⁴. The oligonucleotide fingerprints shown in Fig. 2 clearly show that the recovery of the 5' terminal oligonucleotide (T-13)²⁵ was significantly reduced in conditions of DNA synthesis. Quantitation of the amount of radioactivity present in each of the large oligonucleotides in Fig. 2a and b indicated that approximately 80–90% of the 5' terminal oligonucleotide (T-13) was released during DNA synthesis (Table 1).

The loss of oligonucleotide T-13 observed in fingerprinting experiments suggested that the appearance of discrete oligonucleotides (band A and B of Fig. 1) during DNA synthesis may reflect the action of RNase H on the 5' terminal sequence of ASV RNA. Digestion of the RNA recovered from band A and band B with RNase U₂ (Table 2) yielded the oligonucleotide U₄A, as well as an oligonucleotide with mobility on DEAE-paper identical to that of 7mGpppGomCpCpAp isolated from oligonucleotide T-13. The presence of a capped oligonucleotide in band B was confirmed by digestion of the RNA recovered from band B with alkaline phosphatase and

RNase T₂ (Table 2). The presence of a U₄A and a cap structure in bands A and B, clearly show that these oligonucleotides were derived from the 5' end of the ASV genome. RNase A digestion of band B RNA yielded cap, 1 mol of AU, 1 mol of AC, and no detectable G, suggesting that RNase H cleavage occurs approximately 12–15 nucleotides from the 5' end of the ASV genome.

The data presented here indicate that the avian retrovirus RNA-directed DNA polymerase-associated RNase H activity is capable of specificity removing unique oligonucleotides from the 5' end of the viral genome once DNA transcription has reached the 5' terminus. Previous attempts to demonstrate RNase H activity during reverse transcription were unsuccessful^{22,23,29}. As the efficiency of RNase H activity is sensitive to salt and temperature (data not shown), this failure to detect RNase H activity was probably the result of the sub-optimum reactions used in those studies. In the reaction conditions used here, RNase H activity was observed with both reconstructed RNA:DNA hybrids as well as hybrid structures generated by DNA synthesis. With reconstructed RNA:DNA hybrids, two large oligonucleotides were released by RNase H digestion, whereas during DNA synthesis the shorter of the two oligonucleotides was released along with smaller minor products. Although we cannot explain these differences it is possible that more complete hydrolysis occurs during DNA synthesis. Nevertheless, these studies indicate that RNase H activity results in the release of the terminally redundant RNA sequence when complexed with the complementary nucleotide sequences on tRNA^{trp}-initiated DNA. These findings indicate that long DNA transcripts synthesised from 35S RNA-tRNA^{trp} template-complexes *in vitro*¹⁹ can result from the association of the single-stranded DNA sequence complementary to the 5' terminally repeated sequences present at the 3' end of the viral genome. This interaction could occur by either circularisation of one 35S viral RNA subunit or the end-to-end transcription of two adjacent 35S RNA molecules¹⁴.

We thank S. Kanellos and P. Wilkie for technical assistance and S. Merritt for typing this manuscript. We also thank Drs Bishop, Ostrow, Collins, and Leis for useful discussions for providing data before publication. Drs J. Gruber and J. Beard, and the Office of Program Resources and Logistics, Viral Cancer Program, Division of Cancer Cause and Prevention, NCI kindly provided reagents. This work was supported by NIH research grants CA 18303, CA 21011 (to A.J.F.) and CA 16553 (J.P.) from the NCI.

MARC S. COLLETT*
PETER DIERKS
J. THOMAS PARSONS

Department of Microbiology,
University of Virginia Medical School,
Charlottesville, Virginia 22901

ANTHONY J. FARAST†

Department of Microbiology,
University of Minnesota Medical School,
Minneapolis, Minnesota 55455

Received 25 August; accepted 14 November 1977.

*Present address. Department of Pathology, University of Colorado, Medical Center, Denver, Colorado 802620.

†To whom correspondence should be addressed.

- Haseltine, W. A., Maxam, A. L. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 989–993 (1977).
- Schwartz, D. E., Zamecnik, P. C. & Weith, H. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 994–998 (1977).
- Coffin, J. M. & Haseltine, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1908–1912 (1977).
- Collett, M. S., Dierks, P., Cahill, J. F., Faras, A. J. & Parsons, J. T. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2389–2392 (1977).
- Shine, J., Czernilofsky, A., Friedrich, R., Bishop, J. M. & Goodman, H. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1473–1477 (1977).
- Stoll, E., Billeter, M. A., Palmenberg, A. & Weissmann, C. *Cell* **12**, 57 (1978).
- Collett, M. S. & Faras, A. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1329–1332 (1976).
- Coffin, J. M. *Cancer Res.* **36**, 4282–4288 (1976).
- Taylor, J. M. & Illmensee, R. *J. Virol.* **16**, 553–558 (1975).
- Staskus, K., Collett, M. & Faras, A. J. *Virology* **71**, 162–168 (1976).
- Leis, J., Berkower, I. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **70**, 466–470 (1973).
- Grandgenett, D., Gerard, G. & Green, M. J. *Viol.* **10**, 1136–1142 (1973).
- Verma, I. J. *J. Virol.* **15**, 843–854 (1975).
- Collett, M. S. & Faras, A. J. *Virology* (in the press).
- Collett, M. S. & Faras, A. J. *J. Virol.* **16**, 1220–1228 (1975).

Fig. 2 Release of the 5' terminally-located oligonucleotide T-13 during reverse transcription of ³²P-labelled viral RNA. ³²P-labelled viral RNA (7 × 10⁶ c.p.m.) was incubated with the reverse transcriptase in the presence and absence of deoxynucleoside triphosphates in reaction conditions described in Fig. 1a. The reactions were terminated after 30 min incubation by the addition of EDTA (final 20 mM) and sodium dodecyl sulphate (1%). Samples were heated for 1 min at 100 °C, quenched in ice, and then precipitated with isopropanol in the presence of 1 M NaCl. The samples were centrifuged at 7,000 r.p.m. for 45 min, resuspended in 0.02 M Tris-HCl (pH 7.4), 0.01 M EDTA, treated with T1 RNase, and subjected to two-dimensional electrophoresis as previously described⁴. a, ³²P-RNA incubated with reverse transcriptase in the presence of deoxynucleoside triphosphates. b, ³²P-RNA incubated with reverse transcriptase in the absence of deoxynucleoside triphosphates. Quantitation of the ³²P radioactivity in the numbered oligonucleotides in both panels, was carried out as described previously⁴.

16. Rothenberg, E. & Baltimore, D. *J. Virol.* **17**, 168–174 (1976).
17. Junghans, R. P., Duesberg, P. & Knight, C. A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4895–4899 (1975).
18. Rothenberg, E. & Baltimore, D. *J. Virol.* **21**, 168–178 (1977).
19. Collett, M. S. & Faras, A. J. *J. Virol.* **22**, 86–96 (1977).
20. Haseltine, W. A., Kleid, D. G., Panet, A., Rothenberg, E. & Baltimore, D. *J. molec. Biol.* **106**, 109–131 (1976).
21. Leis, J. *et al. Virology* (in the press).
22. Brewer, L. & Wells, R. *J. Virol.* **14**, 1494–1502 (1974).
23. Collett, M. S. & Faras, A. J. *J. Virol.* **17**, 291–295 (1976).
24. Modak, M. & Marcus, S. *J. Virol.* **22**, 243–246 (1977).
25. Cashion, L. M., Joho, R. H., Planitz, M. A., Billeter, M. A. & Weissman, C. *Nature* **262**, 186–190 (1976).
26. Mölling, K. *et al. Nature new Biol.* **234**, 240–243 (1971).
27. Keller, W. & Crouch, R. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3360–3364 (1972).
28. Grandgenett, D., Gerard, G. & Green, M. *J. Virol.* **10**, 1136–1142 (1973).
29. Darlix, J., Bromley, P. & Spahr, P. *J. Virol.* **22**, 118–129 (1977).
30. Faras, A. J. & Dibble, N. A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 859–863 (1975).
31. Leong, J. *et al. J. Virol.* **9**, 891–902 (1972).
32. Furuichi, Y., Shtatkin, A. J., Starnezer, E. & Bishop, J. M. *Nature* **257**, 618–620 (1975).

Primase initiates Okazaki pieces during polyoma DNA synthesis

SHORT fragments (Okazaki pieces) are formed as intermediates during elongation of nascent DNA strands¹. In isolated nuclei from polyoma-virus infected 3T6 cells Okazaki pieces are initiated by a short RNA primer (initiator RNA, iRNA) with the approximate size of a decanucleotide². The nature of the polymerase responsible for primer synthesis is not known, but high concentrations of α -amanitin, a drug which inhibits both RNA polymerase II (ref. 3) and III (ref. 4), have no effect on primer formation (unpublished results). We now present evidence that iRNA is not synthesised by any of the known RNA polymerases, but that, instead, a mammalian enzyme is involved which in its action corresponds to the *dnaG* gene product of *E. coli*. The bacterial enzyme was named primase⁵. It participates in the initiation of DNA strands of some single-stranded bacteriophages and has been suggested to be involved in the synthesis of bacterial Okazaki pieces^{6,7}. One of the distinguishing features of the *E. coli* enzyme is its ability to use both ribonucleotides and deoxyribonucleotides for primer synthesis^{5,8}.

With isolated cell nuclei, the elongation of polyoma DNA strands required the presence of the four deoxynucleoside triphosphates and ATP (ref. 9). Further addition of CTP, GTP and UTP gave some stimulation which became more obvious when washed nuclei were used¹⁰. The Okazaki pieces formed as intermediates during chain elongation contained iRNA at their 5'-ends. Characterisation of iRNA depended mainly on two types of experiments which involved labelling of the two ends of the molecule (Fig. 1). In the first type 5'-ends were labelled from either β -³²P-ATP or β -³²P-GTP and after digestion with pancreatic DNase, iRNA was characterised by gel electrophoresis⁹. In the second, 3'-ends were labelled from α -³²P deoxynucleoside triphosphates and after alkaline hydrolysis, isotope was transferred from the deoxynucleotide at the RNA-DNA link to the neighbouring ribonucleotide¹¹. Such experiments established that iRNA begins at its 5'-end with either ATP or GTP and contains all four ribonucleotides at its 3'-end. iRNA has no specific base sequence, but shows a pronounced homogeneity in size.

These earlier results were obtained with all four ribonucleotides added during incubation of nuclei. In order to determine how the Okazaki pieces are synthesised when no ribonucleotides are added, washed, polyoma-infected nuclei were incubated with β -³²P-ATP in the absence of the other three ribonucleotides, in otherwise optimal conditions⁹ for polyoma DNA synthesis and replicative intermediates of polyoma DNA were isolated and digested with pancreatic DNase². Any iRNA formed would be released by this treatment and should be labelled with ³²P-ATP at its 5'-end (compare Fig. 1). iRNA was

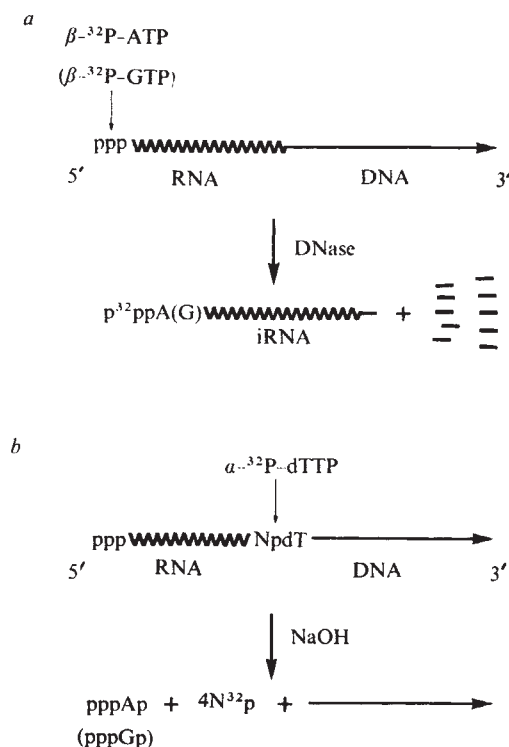


Fig. 1 Experimental design for labelling and characterisation of initiator RNA covalently attached to 5'-ends of nascent DNA strands. *a*, 5'-ends of iRNA labelled with either β -³²P-ATP or β -³²P-GTP. ATP is the preferred starting nucleotide² (2.4 ATP per 1 GTP). Digestion with pancreatic DNase releases iRNA which is purified and characterised by gel electrophoresis⁹. *b*, 3'-ends of iRNA labelled with α -³²P-dTTP or other deoxynucleoside triphosphates. After alkaline hydrolysis ³²P is recovered in the ribonucleotide originally present at the RNA-DNA link¹¹.

prepared in parallel from an identical incubation which, in addition to ³²P-ATP contained CTP, GTP and UTP. Both samples were then analysed by gel electrophoresis with the results shown in Fig. 2*a*. As found earlier², iRNA synthesised in the presence of all four ribonucleotides migrated as a sharp peak at the position of a decanucleotide. In contrast, the ³²P-labelled material synthesised with only ATP, showed a much broader size distribution and the average length of the molecules was smaller. In a second experiment, iRNA was labelled with

Table 1 Transfer of isotope from α -³²P-dTTP to ribomononucleotides and oligonucleotides

Ribonucleoside triphosphates added during incubation	Cp	Mononucleotides			Sum	Dinucleotides	Trinucleotides	Tetranucleotides
		Up	Ap	Gp				
		Transfer $\left(\frac{\text{c.p.m. in fraction}}{\text{c.p.m. in DNA}} \right) \times 1,000$						
ATP, GTP	0.14	0.18	1.25	1.25	2.8	0.81	0.89	0.30
ATP, GTP, UTP, CTP	0.42	0.63	0.52	0.37	1.9	0.13	0.25	0.11

Polyoma DNA was synthesised in washed nuclei as described for experiment *b* in Fig. 2, but with α -³²P-dTTP and non-labelled GTP. Nascent DNA was prepared and hydrolysed with KOH (ref. 11). The acid-soluble hydrolysate was first chromatographed on DEAE-cellulose in 7 M urea¹² to separate mononucleotides and groups of oligonucleotides. Mononucleotides were then further separated on Dowex 1 (ref. 11). The total radioactivity in each fraction was used to calculate the transfer of ³²P from dTTP. For the calculations of 3'-ends, we assumed that the transfer from dTTP amounted to 25% of the total transfer at the RNA-DNA link¹¹.

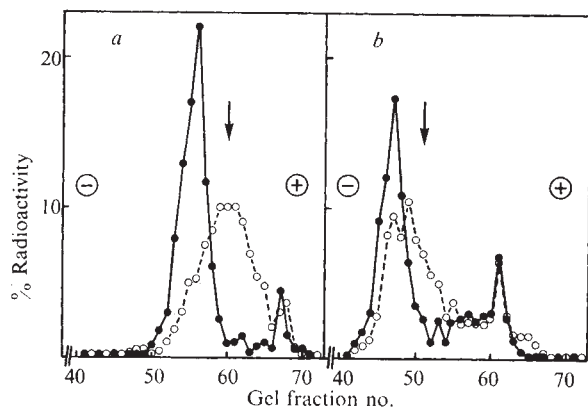


Fig. 2 Gel electrophoresis of iRNA labelled with β - ^{32}P -ATP (a) or β - ^{32}P -GTP (b). In both experiments, iRNA was prepared by DNase treating of replicating polyoma DNA obtained from polyoma infected, washed 3T6 nuclei¹⁰ incubated at 25 °C for 5 min. ●, Experiments carried out in optimal conditions^{9,10} which include the presence of all four ribonucleoside triphosphates, with 0.4 mM β - ^{32}P -ATP in a and 20 μM β - ^{32}P -GTP in b. In each case, the position of the main peak relative to the bromphenol blue marker (arrow) corresponds to that of a decanucleotide, as found previously². ○, Parallel experiments without CTP, UTP and GTP (a) or without CTP and UTP (b). All data plotted as percentage of total radioactivity in each fraction, to allow direct comparison of results.

β - ^{32}P -GTP, with and without added CTP+UTP and analysed by gel electrophoresis (Fig. 2b). Again, deletion of the two pyrimidine ribonucleotides resulted in a broader size distribution and shorter average length of the labelled material.

In these two experiments addition of ribonucleotides increased both total DNA synthesis (measured by incorporation of ^3H -dTTP into polyoma DNA) and primer synthesis (measured by ^{32}P -labelling). In the first experiment, addition of the three ribonucleotides increased DNA synthesis by 39% and primer synthesis by 118%; in the second experiment, the corresponding increases on addition of UTP+CTP were 23 and 66% respectively. Thus, in both cases ribonucleotides had a stronger stimulatory effect on primer synthesis than on DNA synthesis in general.

We next investigated how addition of UTP+CTP influenced the transfer of isotope from α - ^{32}P -dTTP to the 3'-end of iRNA. Deletion of the two ribonucleotides from the incubation mixture decreased DNA synthesis by 18%. In contrast, the total amount of ^{32}P transferred to ribonucleotides after alkaline hydrolysis increased by 18% (Table 1). Deletion of these two nucleotides also caused a considerable change in the relative transfer of ^{32}P to the four ribonucleotides produced by alkaline hydrolysis and the distribution of isotope seemed to reflect directly the presence of a given ribonucleoside triphosphate during the incubation (Table 1).

The increase in total isotope transfer from α - ^{32}P -dTTP caused by deletion of UTP and CTP could hardly be the result of increased primer synthesis with concomitant increase in the number of 3'-ends, since the same deletion resulted in a decrease in the number of 5'-ends. Calculations of the apparent stoichiometries between 3'- and 5'-ends of primer formed in the two experimental conditions give values of 1.18 in the presence of all four ribonucleoside triphosphates and 2.13 in the presence of only ATP+GTP. This apparent excess of 3'-ends could possibly be explained by internal incorporation of α - ^{32}P -dTTP into iRNA. In that case, one would expect that alkaline hydrolysis would result in some transfer of ^{32}P to oligonucleotides as well as to mononucleotides. We therefore chromatographed the material formed by alkaline hydrolysis on DEAE-cellulose in 7 M urea¹² and determined the amount of ^{32}P present in the separated di-, tri- and tetranucleotide fractions (Table 1). In the presence of all four ribonucleotides, small but significant amounts of ^{32}P were recovered in these fractions. Deletion of CTP and UTP increased these values three- to six-fold. In separate experiments

not detailed here, isotope from ^3H -labelled ribonucleoside triphosphates which had been incorporated into the primer could also in part be recovered in the oligonucleotide fraction after alkaline hydrolysis.

From the data in Table 1 we suggest that the primer formed in the absence of UTP and CTP consisted of a mixed ribo-deoxyoligomer. This would also explain why deletion of ribonucleotides resulted in the formation of shorter and polydisperse 'iRNA' after DNase I treatment (Fig. 2), since this enzyme would cleave sufficiently large blocks of internally located deoxynucleotides.

In conclusion, our results indicate that the mammalian primase which catalyses the initiation of Okazaki pieces of polyoma DNA can substitute deoxynucleotides for ribonucleotides and thus resembles the dnaG product of *E. coli*. The mammalian enzyme seems to prefer ribonucleotides for primer formation, but in the presence of equimolar concentrations of ribo- and deoxyribonucleotides, a certain amount of 'misincorporation' of deoxynucleotides can also take place.

This investigation was supported by a grant from the Swedish Cancer Society.

ROLF ELIASSON
PETER REICHARD

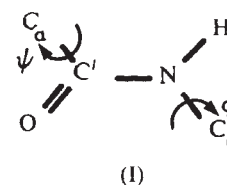
Medical Nobel Institute,
Department of Biochemistry I,
Karolinska Institute,
S-104 01 Stockholm, Sweden

Received 7 November 1977; accepted 23 January 1978.

- Okazaki, R. *et al.* *Cold Spring Harb. Symp. quant Biol.* **33**, 129-143 (1948).
- Reichard, P., Eliasson, R. & Söderman, G. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4901-4905 (1974).
- Kedinger, C., Gniadowski, M., Mandel, J. L., Jr, Gissinger, F. & Chambon, P. *Biochem. biophys. Res. Commun.* **38**, 165-171 (1970).
- Weinmann, R. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1790-1794 (1974).
- Rowen, L. & Kornberg, A. *J. biol. Chem.* (in the press).
- Bouché, J. P., Zechel, K. & Kornberg, A. *J. biol. Chem.* **250**, 5995-6001 (1975).
- Lark, K. G. *Nature new Biol.* **240**, 237-240 (1972).
- Wickner, S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2815-2819 (1977).
- Winnacker, E. -L., Magnusson, G. & Reichard, P. *J. molec. Biol.* **72**, 523-537 (1972).
- Otto, B. & Reichard, P. *J. Virol.* **15**, 259-267 (1975).
- Pigiet, V., Eliasson, R. & Reichard, P. *J. molec. Biol.* **84**, 197-216 (1974).
- Tomlinson, R. V. & Tener, G. M. *J. Am. chem. Soc.* **84**, 2644-2645 (1962).

Regular and non-regular secondary structure in globular proteins

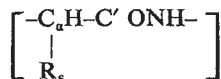
X-RAY crystallographic studies on globular proteins reveal segments whose secondary structure resembles the α -helical and β -extended chain models for polypeptides^{1,2}. These segments are joined by relatively nonregular structures, particularly at bends in the polypeptide backbone. Marked departures from some of the structural criteria laid down in the development of the α and β models occur in these non-regular segments. Here, neither the principle of the constancy of the angles of rotation ψ and ϕ around the C_α - C' and N - C_α bonds respectively³ of the peptide bond (I), nor that of maximisation of hydrogen bond formation between peptide groups applies⁴.



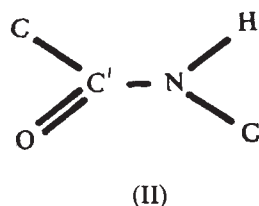
Perturbation of the backbone geometry of a polypeptide chain is generally attributed to short-range noncovalently bonded interactions between side-chain and backbone atoms⁵. The importance of such interactions in determining backbone geometry is clearly shown in poly- α -amino acids and their analogues. The close proximity of a heteroatom or of a bulky non-polar alkyl group to the polypeptide backbone results in its adopting the β -conformation rather than the more usual α -helical form⁶.

Earlier statistical studies on the relationship between non-regular secondary structure and amino acid content⁷ or amino acid sequence⁸ suggest that similar polar and steric factors perturb the backbone geometry of globular proteins. More recent statistical studies (refs 9–14 and M. Levitt, personal communication) on the relationship between amino acid sequence and backbone conformation indicate that nonregular secondary structure may be induced by the presence of proline and glycine, and also by another group of amino acid residues, such as asparagine, aspartic acid and serine, whose side chains contain an oxygen atom close to the protein backbone. The possibility that steric factors contribute to the induction of nonregular secondary structure has also been suggested¹³.

In a globular protein the steric and polar properties of the side chain (R_s) vary from one peptide unit



to the next along the polypeptide chain. It is conceivable that a combination of these side-chain properties could be responsible for the perturbation both of the geometry of the backbone, and of the formation of hydrogen bonds between peptide groups. Maximisation of short range noncovalently bonded interactions could lead to an alternative stable hydrogen-bonded form of the peptide bond. Further, if this supposition is valid, a second, stable hydrogen-bonded form of the analogous secondary amide bond (II) should be observable. The suffix α is omitted from (II)



compared with (I) because specific stereochemical configurations at the carbonyl and amido carbon atoms (C_α) were not taken into account in the model compounds examined.

In *N*-alkylamides of the type $R.C'ONH.R'$ the perturbatory effect of changes in non-covalently bonded interactions may be maximised by bringing polar atoms and bulky groups as close as possible to the amide link. The steric and polar properties of R and R' were varied systematically and the effect of these variations on the conformational properties of the amide bond (II) was examined experimentally using infrared spectroscopic techniques¹⁵.

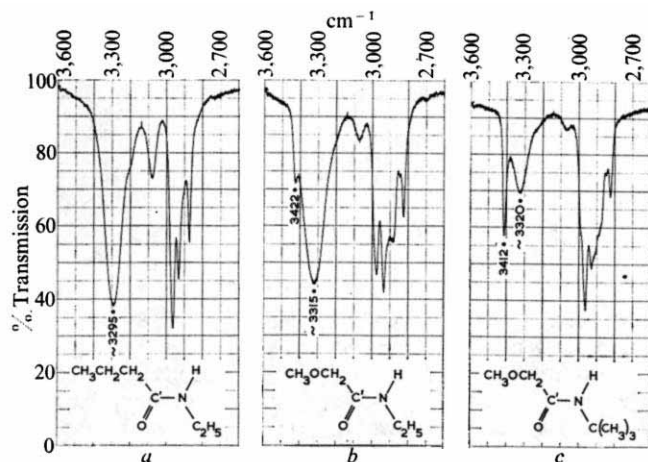
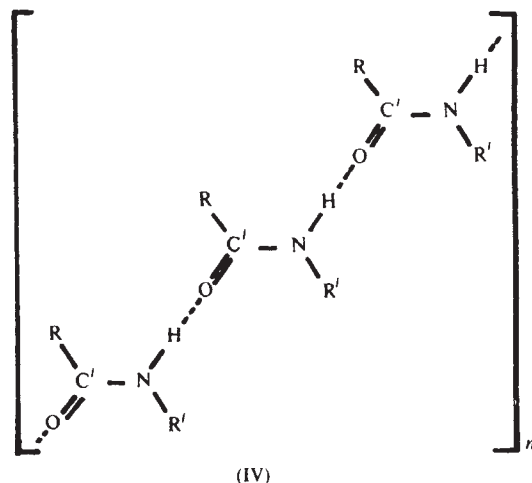
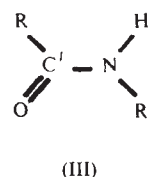


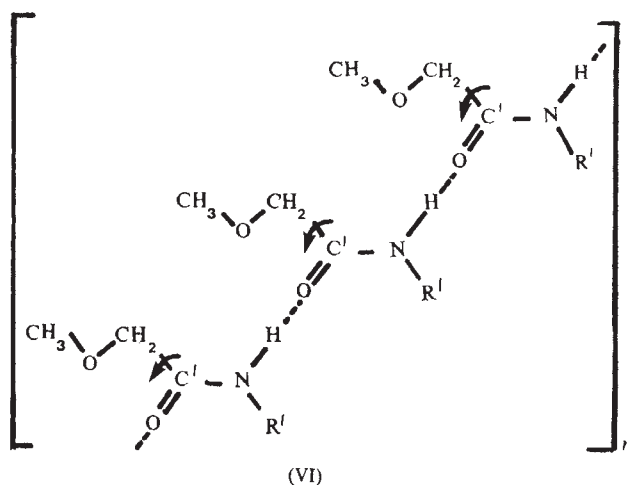
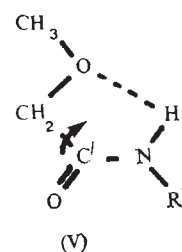
Fig. 1 3 μ m spectra of liquid capillary films of *a*, *N*-ethylbutyramide; *b*, *N*-ethylmethoxyacetamide; and *c*, *N*-*tert*-butylmethoxyacetamide.

N-alkylamides are predominantly in the *trans*, planar conformation (III) (ref. 16). In the liquid state they are normally associated, as a result of interamide hydrogen bonding, into multimeric species (IV).



The multimers give rise to a broad, intense, associated NH stretching absorption band near 3,300 cm^{-1} in the infrared region of the spectrum (Fig. 1a)^{16,17}.

Some analogous *N*-alkylfluoroacetamides and *N*-alkylmethoxyacetamides also show a broad associated NH stretching band near 3,300 cm^{-1} due to multimeric species (VI). But in this instance there is an additional and unexpected sharp band between 3,410 cm^{-1} and 3,440 cm^{-1} (refs 15, 18).

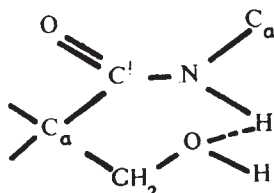


The sharp bands at $3,422\text{ cm}^{-1}$ and $3,412\text{ cm}^{-1}$ in, for example, the spectra of liquid *N*-ethylmethoxyacetamide (Fig. 1b) and *N*-*tert*-butylmethoxyacetamide (Fig. 1c) are assigned to the NH stretching mode of intramolecularly hydrogen-bonded isomers (V). These assignments are made by analogy with the comparable position of the single, sharp absorptions observed at $3,435\text{ cm}^{-1}$ and $3,416\text{ cm}^{-1}$, respectively, in 0.001 M carbon tetrachloride solution spectra of the same compounds¹⁵. In these conditions, the corresponding unmodified *N*-alkylamides ($\text{CH}_3\text{CH}_2\text{CH}_2\text{C}'\text{ONH.R}'$) absorb at $3,456\text{ cm}^{-1}$ and $3,445\text{ cm}^{-1}$. At the high dilutions used, most of the molecules are in the monomeric state, and the small shifts to lower frequencies (21 cm^{-1} and 29 cm^{-1} , respectively) in the *N*-alkylmethoxyacetamides can be attributed to their being stabilised by an intramolecular hydrogen bond.

If the reasonable assumption is made that the intrinsic intensities of the NH stretching bands in the liquid state spectra do not vary appreciably as the nonpolar *N*-alkyl substituent is varied, the contribution of the intramolecularly hydrogen-bonded species to the equilibrium increases as the size of the *N*-alkyl group (R') is increased, from that of an ethyl group to that of a *tert*-butyl group. Simultaneously with the formation of an intramolecular hydrogen bond, the angle of rotation around the $\text{CH}_2\text{-C}'$ bond is altered as shown schematically in (V) and (VI).

New equilibria are observed in the model compounds when the perturbatory influence of polar atoms and nonpolar groups on either side of a single amide bond is maximised. The primary condition for the co-existence of conformationally isomeric forms in liquid *N*-alkylmethoxyacetamides is the proximity of an electronegative oxygen atom to the amido hydrogen atom (V). Consequently, the amido hydrogens in some of the molecules form intramolecular hydrogen bonds, and the angle of rotation around the $\text{CH}_2\text{-C}'$ bond in these molecules is altered.

An analogous short range interaction between a side-chain oxygen and a backbone peptide hydrogen has been shown¹⁹ to be possible in the case of, for example, serine (VII) when this residue is in a non-helical conformation.



(VII)

Thus, in a globular protein, a side-chain oxygen atom close to the backbone could compete for, or perturb, a backbone peptide hydrogen (whose usual function is to form an interpeptide hydrogen bond) and simultaneously twist the $\text{C}_\alpha\text{-C}'$ bond, thus altering the angle ψ (I). Two of the structural constraints implicit in regular α and β models for secondary structure could be modified simultaneously: first, the maximisation of hydrogen-bond formation between peptide groups, and second, the constancy of the angle of rotation around the $\text{C}_\alpha\text{-C}'$ bond.

These new structural considerations suggest a mechanism whereby non-regularity is induced in the secondary structure of those segments of globular proteins with amino acid sequences containing an above average concentration of side chains with an oxygen atom close to the backbone. Furthermore, the suggested mechanism is consistent with a feature of amino acid frequency diagrams plotted by Scheraga *et al.*²⁰. The diagrams consist of plots of the angles of rotation ψ and ϕ (defined in (I) above) for all the amino acid residues in eight proteins. The data are plotted on separate $\phi\text{-}\psi$ conformational maps for each of the twenty amino acids normally found in globular proteins. With the exception of glycine, the points

tend to cluster in two areas representing α -helical and β -extended chain conformations. But, as those authors point out, the area between these highly populated areas, which is normally forbidden, becomes significantly populated in the direction of the ψ -axis in the case of serine, threonine and asparagine.

The direction of change of the equilibria observed in the *N*-alkylmethoxyacetamides illustrates the fundamental role of the probability distribution of labile amide hydrogens in determining the position of the equilibrium for each individual compound (Fig. 1). The carbonyl and methoxy oxygen atoms act as alternative proton-accepting sites. As the size of the nonpolar substituent at the amido nitrogen is increased, hydrogen bonding to the methoxy oxygen becomes more probable, and so a conformational change takes place in an increasing proportion of the molecules.

From a dynamic point of view, the relative proportion of the two conformationally isomeric forms observed in a given compound is determined by the relative proportion of its lifetime which a tautomeric proton spends in the vicinity of a carbonyl oxygen or a methoxy oxygen atom. This, in turn, is dependent on the steric and polar factors operating in the neighbourhood of a given amido hydrogen atom at a given instant.

The equilibria illustrate limiting cases: their existence provides a starting point for discussion of the conformational properties of globular proteins in dynamic terms, and for the introduction of time as a factor in the perturbatory influence of side chain oxygen atoms on backbone conformation. The observed equilibria and their direction of change suggest that, as a consequence of very rapid variations in the steric and polar environment of a peptide hydrogen, conformational flexibility is induced in the protein backbone.

From this dynamic point of view, the occurrence of bends or nonregular segments where there are higher-than-average concentrations of, for example, asparagine, aspartic acid, and serine, could be attributed to the labile peptide hydrogens spending, on average, a larger fraction of their lifetime in the vicinity of side-chain oxygen atoms close to the backbone, compared with peptide hydrogens in regular α or β segments, which would spend a larger fraction of their lifetime in the vicinity of backbone carbonyl oxygen atoms. Thus the probability distribution of labile backbone peptide hydrogens at a given instant is likely to be a fundamental factor determining the overall backbone conformation of a native globular protein molecule at that instant of time.

I thank Professor W. V. Mayneord, John Crookall-Greening, and Dr Ian Gilbert for constructive criticism of the manuscript.

R. LUMLEY JONES*

Chester Beatty Research Institute,
Institute of Cancer Research, Royal Cancer Hospital,
Fulham Road, London SW3, UK

Received 27 May 1977; accepted 13 January 1978.

*Present address: Department of Chemical Pathology, St Bartholomew's Hospital, 51-53 Bartholomew Close, London EC1, UK.

- Pauling, L. & Corey, R. B. *Proc. natn. Acad. Sci. U.S.A.* **37**, 205-211 (1951).
- Pauling, L., Corey, R. B. & Branson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **37**, 251-256 (1951).
- J. molec. Biol.* **52**, 1-17 (1970).
- Edsall, J. T. in *Structural Chemistry and Molecular Biology* (eds Rich, A. & Davidson, N.) 88-97 (W. H. Freeman, San Francisco and London, 1968).
- Kotelchuck, D. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **61**, 1163-1170 (1968).
- Blout, E. R., de Lozé, C., Bloom, S. M. & Fasman, G. D. *J. Am. chem. Soc.* **82**, 3787-3789 (1960).
- Davies, D. R. *J. molec. Biol.* **9**, 605-609 (1964).
- Pittsytyn, O. B. *J. molec. Biol.* **42**, 501-510 (1969).
- Lewis, P. N., Momany, F. A. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2293-2297 (1971).
- Kuntz, I. D. *J. Am. chem. Soc.* **94**, 4009-4012 (1972).
- Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 211-221 (1974).
- Robson, B. & Pain, R. H. *Biochem. J.* **141**, 899-904 (1974).
- Robson, B. & Pain, R. H. *Biochem. J.* **141**, 883-897 (1974).
- Finkelstein, A. V., Pittsytyn, O. B. & Kozitsyn, S. A. *Biopolymers* **16**, 497-524 (1977).
- Jones, R. L. *Spectrochim. Acta* **30A**, 1329-1333 (1974).
- Mizushima, S. *et al. J. Am. chem. Soc.* **72**, 3490-3494 (1950).
- Beer, M., Kessler, H. B. & Sutherland, G. B. M. *J. chem. Phys.* **29**, 1097-1104 (1958).
- Letaw, H., Jr & Gropp, A. H. *J. chem. Phys.* **21**, 1621-1627 (1953).
- Lewis, P. N., Momany, F. A. & Scheraga, H. A. *Israel J. Chem.* **11**, 121-152 (1973).
- Burgess, A. W., Ponnuswamy, P. K. & Scheraga, H. A. *Israel J. Chem.* **12**, 239-286 (1974).

matters arising

'Unmixing' of the deep-sea record and the deglacial meltwater spike

BERGER, Johnson and Killingley¹ have suggested that at about 12,000 yr b.p. a large part of the world ocean was covered by a layer of meltwater from the Laurentide ice sheets. Such a finding is of interest in that it helps to explain the relatively rapid melting of the Würm/Wisconsin glacial ice sheets. Mason² calculates that the total excess of insolation due to the Milankovitch mechanism³ north of 45°N between 16,000 and 6,000 yr b.p., when about 90% of the ice melted, was 1.0×10^{25} cal, while about 3.2×10^{24} cal was required to supply the latent heat of fusion of the ice sheets. The initial melting of the ice sheets was probably rapid and this gives rise to problems of energy supply, since little solar insolation is absorbed directly by the high albedo surfaces of ice sheets. Saltzman⁴ and Newell⁵ have suggested that mean global ocean temperatures increased during the period of growth of ice coverage, reaching a maximum sometime after the time of maximum ice extent and a minimum sometime after the interglacial. The latent heat of fusion for melting the ice sheets could then be supplied by the deep oceans. This result is not confirmed by ocean-surface temperature estimates for 18,000 yr b.p. by CLIMAP project members⁶, though changes in deep ocean temperatures are very uncertain. Adam⁷ has pointed out the importance of a layer of meltwater on ocean-surface temperature. The stratification of salt in such a layer will produce a stable stratification. Further, solar insolation warms the top, adding a thermal density stratification to the already strong salinity density stratification. Both factors inhibit the downward mixing of absorbed solar heat, raising the surface temperature of the ocean in the presence of a surface meltwater layer above the value that would be found without one. This effect would be at a maximum during the summer when the Northern Hemisphere was receiving extra insolation due to the Milankovitch mechanism. Seawater has a low albedo (10% or less) so it would be a very effective absorber of the extra heat. Very rapid summer melting of the ice-sheets could therefore be possible using the extra heat supplied by the Milankovitch mechanism, which was absorbed

by the meltwater on the ocean surface and then transferred to the ice sheets by atmospheric advection. Further later warming could be caused by the release of carbon dioxide to the atmosphere as suggested by Berger, Johnson and Killingley.

J. G. LOCKWOOD

*School of Geography,
University of Leeds,
Leeds, UK*

1. Berger, W. H., Johnson, R. F. & Killingley, J. S. *Nature* 269, 661-663 (1977).
2. Mason, B. J. *Q. Jl R. Met. Soc.* 102, 473-498 (1976).
3. Vernekar, A. D. *Met. Mon.* 12, No. 34 (1972).
4. Saltzman, B. *Tellus* 29, 205-212 (1977).
5. Newell, R. E. *Quart. Res.* 4, 117-127 (1974).
6. *Science* 191, 1131-1137 (1976).
7. Adam, D. P. *Quart. Res.* 5, 161-171 (1975).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words, and the briefest of replies, to the effect that a point is taken, should be considered.

Rare male mating advantage

A SEARCH for and establishment of any frequency dependent component of Darwinian fitness is often likely to meet frustratingly diverse alternatives of interpretation. O'Donald¹ has demonstrated frequency dependence with a model based on a constant female 'preference' parameter for some of our data⁴⁻⁷ on rare male mating advantage. This 'pseudofrequency dependence' ensues in 'multiple choice' matings when the input ratio of type *a*: type *b* males is varied, because as either type becomes rare, its success occurs in greater proportion relative to that of the more common type. The principle that frequency dependency may arise with

application of constant fitness components has been pointed out particularly by Prout² and was previously mentioned by Spiess and Langer³. Thus according to O'Donald, there is no need to invoke a change in a female's preference (we prefer to use the term bias) towards the types of courting males. In effect, the burden of proof for a behavioural interpretation of our data⁴⁻⁷ cited by O'Donald falls on us. We respond here with some evidence in favour of our behavioural interpretation, both published and unpublished.

We have had no evidence in the experiments cited or in subsequent experiments that females show preference toward either type of male when tested at control frequencies (*A* males equal to *B* males). We cannot rule out the possibility that a small fraction of females displays some negligible amount of preference. Our mating data can be explained for controls, as well as most of the cases when female ratios were varied⁵, merely on the basis of relative courtship and mating ability of the male types, not invoking assortative, or preferential parameters among females. Briefly, our hypothesis was as follows: a virgin female at the start of a mating test samples courtship cues from males and becomes habituated to the majority type of male so that she tends to accept a male with a different cue merely because it interrupts the habituation. While O'Donald's model does not include a differential preference among females, we assume that, if preferences were real, they would be likely to differ among types of females and some evidence for assortative mating would be expected. No general evidence for assortative mating was found^{5,6} for AR: PP *Drosophila pseudoobscura* or for Hu: WW *D. persimilis*.

O'Donald's model assumes both types of female to have identical proportions of constant preferences. If so, then varying the ratio of AR:PP among females, for example, should not affect the amount of frequency dependent advantage for males. However, in AR:PP experiments, when ratios of females as well as males were varied simultaneously (Table 4 and Fig. 4 in ref. 5), frequency dependent advantage was much less than when ratios of males alone were varied. No minority advantage occurred for females when their ratios were varied. O'Donald would need to postulate that two female karyotypes have different

preferences in order to conform to such a result, and his model would need to include components for each type of female. Our explanation based on relative mating receptivities for females and courtship activities for males seems simpler and less hypothetical.

O'Donald¹ cited a paper in press⁸, not available to us, where possible behavioural events were proposed which should follow from his preference model. He predicted that "The proportion of preferential matings would therefore decline with increasing density", so that in the double chamber experiments⁷ with a minority ratio in the upper chamber, when extra males were added to the lower chamber, the minority advantage would disappear. O'Donald did not discuss a prediction for the opposite result, namely the induction of a minority advantage with double chamber technique by initiating the experiments with equal numbers of types in the upper chamber, which Ehrman⁹ demonstrated later. In both cases, addition of flies to the lower chamber increased density.

Furthermore, O'Donald considered "A similar result would follow if different male genotypes determined different levels of courtship activity . . . A less responsive female would mate readily with one or more active males and would therefore tend to show a preference that decreased at higher levels of stimulation". These assumptions are not borne out by recent experiments of Spiess and Schwer¹⁰, which have demonstrated that, in spite of courtship intensity differences between two genotypes of male *D. melanogaster*, females displayed no bias toward either type at control frequencies, whether flies were tested as groups of 20 pairs or as single virgin females confined with six males. But when two types of males were introduced at unequal frequencies, the females accepted minority males without alteration in the average time to acceptance; nor did minority males court with any change in their intensity. In fact, Spiess and Schwer¹⁰ found by analysing the sequence of individual courtships that females are less inclined to mate with the type of male which courts them first; Ehrman's observations agree. We have interpreted these results as lending support to the behavioural hypothesis⁴ that a female samples males and their courtship cues, becoming habituated to the majority (or alternatively avoiding the cue of the first male to court) but accepting a male with cue different from that she first detected. Thus behavioural tests lend no evidence for a female preference parameter.

The establishment of either argument as the mechanism to account for rare male mating advantage can only follow more extensive empirical behavioural testing¹⁰⁻¹².

A more substantive proof for a behavioural basis of the male minority

advantage in mating will be published elsewhere by E.B.S.

ELLIOT B. SPIESS

*Department of Biological Sciences,
University of Illinois at Chicago Circle,
Chicago, Illinois 60680*

LEE EHRLMAN

*Division of Natural Sciences,
State University of New York,
Purchase, New York 10577*

1. O'Donald, P. *Nature* 267, 151 (1977).
2. Prout, T. *Genetics* 63, 949 (1969).
3. Spiess, E. B. & Langer, B. *Evolution* 18, 430 (1964).
4. Ehrman, L. & Spiess, E. B. *Am. Nat.* 103, 675 (1969).
5. Spiess, E. B. *Am. Nat.* 102, 363 (1968).
6. Spiess, L. D. & Spiess, E. B. *Am. Nat.* 103, 155 (1969).
7. Ehrman, L. *Am. Nat.* 101, 415 (1967).
8. O'Donald, P. *J. Theor. Pop. Biol.* (in the press).
9. Ehrman, L. *Science* 167, 905 (1970).
10. Spiess, E. B. & Schwer, W. A. *Behav. Genet.* 8, 155 (1978).
11. Leonard, J. & Ehrman, L. *Science* 193, 693 (1976).
12. Ehrman, L. & Parsons, P. *Genetics of Behaviour* Ch. 8 (Sinauer Associates, Sunderland, Massachusetts, 1976).

O'DONALD REPLIES—In restating their own behavioural interpretation of the rare-male advantage in *Drosophila*, Spiess and Ehrman argue that because no evidence of preferential mating has been found at 'control frequencies' when genotypic frequencies are equal, therefore only a small fraction of females could display no more than "some negligible amount of preference". This conclusion does not follow, however. If equal proportions of females prefer two genotypes, then in my model any magnitude of such preferences will give equal expectations of numbers of matings at the control frequencies. In the data, deviations from equal expectations occur at the control frequencies. These deviations, although not significant by themselves in the small sample sizes at these frequencies, are nevertheless much reduced after fitting of the preferences estimated from the data obtained at all frequencies. This is shown in the expectations of the matings in Table 1a of ref. 1.

It is certainly true, as Spiess and Ehrman state, that in fitting their data I assumed equal expression of preference among the female genotypes. If preferences were expressed at different levels in different genotypes, this would produce assortative mating which has not been observed. Spiess and Ehrman's own explanation, which, they say, seems simpler and less hypothetical also depends on the same assumption. Both explanations would become less simple if the assumption were removed from the models.

Spiess and Ehrman then refer to some data of Spiess and Schwer, still in the press. It seems to me that these data clearly refute Spiess and Ehrman's own model, but not the model of female

preference. If males differ in the intensity of their courtship and females become habituated to the courtship cues provided by the males, then at the control frequencies we should expect a bias in favour of the males with the less intense courtship giving fewer cues to which the females would be less likely to become habituated. The absence of bias favouring one or other type of male therefore refutes both my own model based on differences in courtship intensities and Spiess and Ehrman's model, leaving the model of female preference. If females have preferences for one or other of the males, regardless of differences in courtship intensity, then no bias should be found in the matings at the control frequencies.

On evolutionary grounds, Spiess and Ehrman's theory is implausible: a general tendency to mate with the more unusual phenotypes would be maladaptive.

PETER O'DONALD

*Department of Genetics,
University of Cambridge, UK*

Population dynamics and the length of food chains

PIMM and Lawton¹ have claimed that increasing the number of trophic levels tends to increase the time taken for a stable ecosystem to return to equilibrium after a perturbation, and they have suggested that this is the chief reason why food chains are typically short. I show here that the first of these claims is not necessarily true, and that even if it were, the second does not follow from it.

Pimm and Lawton based their conclusions on investigations of the eigenvalues of matrices whose elements were chosen randomly subject to certain constraints on magnitude and sign. Their key finding was that as the number of trophic levels increased, so did the reciprocal of the largest eigenvalue. In fact it is quite easy to see how this came about. In all the models that were considered, of the diagonal elements only those associated with the primary producers were allowed to be non-zero, and these were all chosen within the interval (0, -1). The expected value of the trace of the stability matrix was thus proportional to the number of primary producers. Since the trace of a matrix is equal to the sum of the eigenvalues, and since the probability that at least one eigenvalue of a 4×4 stable matrix will have a real part greater than a certain fixed value clearly decreases as the sum of the eigenvalues decreases, we would expect that the six models would fall into three classes and that in order of increasing stability (in the sense of Pimm and Lawton) these would be (a, b), (c, d, f), (e). This is precisely what was found.

The above argument of course ignores the effect of the off-diagonal terms, but it has been shown² that on average these have little or no effect on the damping; they are chiefly involved in producing oscillations or destabilising the system altogether.

The important factor, therefore, is the total amount of 'damping power' (in the form of self limitation of species) available in the system. The claimed result is valid only if all species other than primary producers are limited solely by interspecific competition or predation, and even then what matters is the ratio of primary producers to other species, and not the number of trophic levels.

Even if the result concerning the number of trophic levels were true there is a further difficulty. Pimm and Lawton have chosen to ignore all those systems which are not locally stable. This is standard practice and indeed gives the correct sample space if we suppose that unstable systems do not survive long enough to be observed³. But if we are to be consistent we must apply the same argument to those systems which are unstable on account of insufficient damping. If we are prepared to accept that complex ecosystems can exist because they have evolved in such a way as to possess the (*a priori* unlikely) property of local stability, then in exactly the same way we would expect to see systems with many trophic levels, even if that property too seems to imply a lower probability of survival.

One can imagine how such a system could evolve. If the only condition we impose on a new species being added to a stable ecosystem is that it neither destroys the local stability nor reduces the damping rate, then there is always at least one obvious niche available: a newcomer should be able to survive providing that it is self limiting and preys on what was previously a top predator but kills sufficiently few of them to leave the previously existing equilibrium effectively unaltered. In this way the number of trophic levels can always be increased. If, however, we take considerations such as energy requirements and body size into account, then this niche may not exist, so such factors seem much better candidates as explanations for the limit on the number of trophic levels.

Why food chains should be as short as they are is still not entirely clear, but it is very unlikely that the answer lies in population dynamics.

P. T. SAUNDERS

Department of Mathematics,
Queen Elizabeth College,
London W8, UK

LAWTON AND PIMM REPLY—We are pleased to see that Saunders¹ shows analytically what we found by simulation². We are unconvinced by his subsequent arguments.

We will deal first with the evolution of "a newcomer . . . that is self limiting and preys on what was previously a top predator but kills sufficiently few of them to leave the previously existing equilibrium effectively unaltered". Such an animal is easy to create mathematically, but not biologically. To behave as suggested the animal will usually have to evolve some system, a territory or peck-order, to provide the self-limitation; but more importantly the self-limitation must also ensure that virtually all the resource (prey) being exploited remains unutilised. This is exactly what Wynne-Edwards envisaged³ and it is generally agreed that it requires group selection to achieve it⁴. Nor is it a general feature of most real food webs⁵⁻⁹.

Even if we are generous, and assume that it can happen, sometimes, then we would predict that Saunders' special food-chains would be longer in more productive ecosystems. This is where we came in²; they are not.

However, there are special situations which can generate ineffective predators without invoking the pitfalls of group-selection. Sea birds cropping some trivial fraction of a fish population (because they cannot dive more than a few metres into the water) provide a reasonable example. Interestingly those food chains which cross the sea-land interface may characteristically be longer than most¹⁰, which is exactly what we would predict and lends support to our hypothesis that in general, food chains are limited by population dynamics and not by energetics.

Saunders' second main point, that "what matters is the ratio of (the number of species of) primary producers to other species, and not the number of trophic levels" is more interesting. It is only a counter to our general, and mainly qualitative insight, if the number of independent, self-limited species at the base of food-chains differs markedly from one situation to another. We doubt whether this is so. Most insect herbivores, for example, are monophagous or oligophagous, giving rise to relatively discrete food chains even in species-rich plant communities¹¹⁻¹²; hence in practice the problem may be more apparent than real. It disappears altogether, at least mathematically, if we treat all the food-species taken by polyphagous grazers as constituting one 'population'. This is more or less the same as assuming that the self-limited plant species also show strong interspecific competition. If the number of self-limited species has an upper bound set by interspecific competition (which does not seem unreasonable¹³) then Saunders' result provides an

analytical proof for our simulation results: the more trophic levels per self-limited species, the less likely the system is to be biologically feasible.

Finally, Saunders has directed himself only to our differential equation models. Difference equations, implying different life histories, yield similar qualitative insights although for slightly different reasons. In other words, whilst energetics conspicuously fails to explain why food chains are usually short despite enormous variations in productivity, biologically reasonable population-dynamic models converge to provide a plausible alternative hypothesis.

Our original letter contains three errors, for which we apologise. They are as follows. (1) Column one; the second equation should read

$$\frac{\partial F_j}{\partial X_i} \quad j=1, n \\ i=1, n$$

(2) Figure 1; the legend should read 'down each column has . . . on the species in each row . . .'. (3) Figure 1e; all the positive elements down the main diagonal should be negative.

We thank John Beddington for helpful discussion.

J. H. LAWTON

Department of Biology,
University of York,
Heslington, York, UK

S. L. PIMM

Department of Biological Sciences,
Texas Tech University,
Box 4149, Lubbock,
Texas 79409

1. Saunders, P. T. & Bazin, M. J. *Nature* 256, 102-121 (1975).
2. Pimm, S. L. & Lawton, J. H. *Nature* 268, 329-331 (1977).
3. Wynne-Edwards, V. C. *Animal Dispersion in Relation to Social Behaviour* (Oliver and Boyd, Edinburgh, 1972).
4. Lack, D. *Population Studies of Birds* (Clarendon, Oxford, 1966).
5. Huffaker, C. B. (ed.) *Biological Control* (Plenum, New York, 1974).
6. Mech, L. D. & Karns, P. D. *U.S. For. Serv. Res. Rep. NC-148* (1977).
7. Estes, J. A. & Palmisano, J. F. *Science* 185, 1058-1060 (1974).
8. Paine, R. T. *Am. Nat.* 100, 65-75 (1966).
9. Brooks, J. L. & Dodson, S. I. *Science* 150, 28-35 (1965).
10. Elton, C. S. *Animal Ecology* (Sidgwick and Jackson, London, 1927).
11. Shure, D. J. *Ecol. Monogr.* 43, 1-19 (1973).
12. Wiegert, R. G., Odum, E. P. & Schnell, J. H. *Ecology* 48, 75-83 (1967).
13. May, R. M. (ed.) *Theoretical Ecology* (Blackwell, Oxford, 1976).

Nature Indexes

The Nature index for 1977 is available at £2.50 (UK), US\$5.00 (Rest of World). See page xii for details.

1. Pimm, S. L. & Lawton, J. H. *Nature* 268, 329-331 (1977).
2. Saunders, P. T. & Bazin, M. J. *Nature*, 256, 102-121 (1975).
3. May, R. M. *Math. Biosci.* 12, 59-79 (1971).

reviews

Fungi and the life of man

M. J. Carlile

Fungi, Man and his Environment. By R. C. Cook. Pp. 144. (Longman: London, 1977.) £6.75.

There are several good textbooks on the systematic, physiological and genetic aspects of the fungi. The present work is complementary to these, and provides a readable account of the ways in which fungi directly or indirectly affect the life of man. This is a particularly laudable endeavour, as the relevant literature is widely scattered, often in journals little read by and relatively inaccessible to most biologists. The author stresses that he has not attempted to produce a textbook or a uniform coverage of the field, but has written about topics that he has found interesting. The most important omissions, as the author admits, are the role of fungi in the preparation of traditional foods and drinks, and in the production of antibiotics. Apart from this, most of the economically important aspects of the fungi are covered.

The first chapter is a brief introduction to the fungi for non-mycologists. The next, "nutrients and niches", deals with the habitats of fungi in relation to nutrition and physiology, and the related topic of biodegradation, both beneficial, as in composts, and harmful, as in the various forms of biodeterioration such as timber decay and the spoilage of aviation fuel, food and stored products.

Plant disease is then discussed (chapter 3), with emphasis on recent catastrophic episodes such as Dutch Elm Disease in the United Kingdom and Jarrah Dieback in Australia. A chapter on symbiosis follows, with stress on lichens as indicators of atmospheric pollution and mycorrhiza as a major factor in the nutrition of trees and perhaps crop plants. Chapter 5 discusses fungal diseases of man and animals, including fish, and is followed by an account (chapter 6) of biological control, by fungi and of fungi. Chapter 7 is concerned with fungus-insect relationships: fungus gardens of ants, fungi as insect endosymbionts, and even insects as slaves of the fungi—scale insects extracting nutrients from plants on behalf of their fungus captors.

The chapter (8) on mycotoxins concentrates on the gruesome consequences of consuming one of the few really deadly mushrooms, *Amanita*

phalloides, the effects of ergot and its derivatives, and the aflatoxin story. There is a short chapter (9) on the industrial production of microbial protein, and a longer one (10) on fungi and magic: fairy rings, luminous wood and hallucinogens. The book ends with a short bibliography of key references for each chapter, and an index.

There are a few errors, almost inevitable in a book covering so wide a field, but none that seriously detract from the value of the work. It is to be hoped that the publishers will refrain from continued use of such small print, which I found liable to cause eyestrain

except under very good light. The price seems rather high for such a slim volume, but is perhaps explained by the copious photographs and drawings, which are attractive and instructive, enhancing the value of the text. The author and publisher are to be congratulated on a pioneering work, which admirably demonstrates the practical importance of fungi, and which should be of value to both lecturers and students. □

M. J. Carlile is Senior Lecturer in Microbiology in the Department of Biochemistry, Imperial College, London, UK.

Aspects of high energy physics

Deeper Pathways in High-Energy Physics. Edited by A. Perlmutter and L. F. Scott. Pp. 743. (Plenum: New York, 1977.) \$71.40.

THIS volume contains papers presented at the fourteenth annual Coral Gables Conference. The conference opened with a session on "Polarized Particles and Spin Effects in High Energy Physics". The recent development of a polarised proton beam at Argonne National Laboratory makes it possible to measure new properties of reaction amplitudes and the new data reported here are a challenge to all models of hadron reactions. The novel questions posed by such data are discussed in a useful review by Halzen.

The next session on "Supersymmetry and Supergravity" was devoted to important work on local supersymmetries carried out during 1976. One attraction of supersymmetries is that they link together particles of different spins in a single multiplet, in which there is a connection between the spin content and the internal symmetry content. However, despite the large amount of work which has been done on these mathematically elegant models "there are as yet no signals from experiment that nature is aware of our efforts, (so) we look for theoretical signals", to quote the article by Freedman. The possibility that the theory may be renormalisable is perhaps the most interesting signal. This is discussed in a particularly clear article by van Nieuwenhuizen and Grisaru. In supergravity

(which only differs from Einstein's theory at short distances), the contributions to amplitudes due to exchanges of different particles in a multiplet are connected and tend to cancel in a remarkable way. This renders one- and two-loop amplitudes finite in certain models. If amplitudes turn out to be finite to all orders, supergravity would be the first renormalisable quantum theory of gravity, with which we could calculate to all orders in principle. Whether this happens and whether the models have anything to do with the real world are still open questions.

The remaining sessions were mainly on questions connected with what (unfortunately) is coming to be known as QFD—quantum flavourdynamics (the theory of weak-electromagnetic interactions, which couple to 'flavour' quantum numbers—charge, strangeness, charm, and so on) and QCD—quantum chromodynamics (a model in which strong interactions are generated by fields coupled to the 'colour' quantum number of hadrons). Despite changes in the data during the past year, the review of the status of QFD by Fritzsch and the discussion of how to search for heavy particles by Barnett are still valuable. The same is true of the paper by Cheng and Li on the decay $\mu \rightarrow e\gamma$, although rumours that this decay had been observed turned out to be false. There is also a useful review by Bouchiat of the atomic physics experiments, which can constrain neutral couplings so strongly, and a good summary of ideas about charmed particle decays by Rosner.

In contrast to QFD, where the gauge principle seems established and where we know how to calculate but the specific choice of model is open, there is only one plausible model in QCD, but too little is known yet about how to solve it to judge whether it is correct. Progress in this developing field is described in papers by Gross, Cornwall and 't Hooft. A paper by Politzer describes some work on the relatively well established applications of QCD to deep inelastic processes. There is also a useful review by Lipkin of the empirical status of the so-called Zweig rule, which is supposed to inhibit the decay of the J/ψ to hadrons so strongly. It is believed that QCD may explain this rule but in truth it is hardly understood at all. Similarly, it is hoped that QCD might account for the successes of hard scattering models of large P_T hadronic processes of the sort described in the paper by Feynman.

Many of the contributions to this volume are more or less directly based

on, or form the basis for, papers which have now been published in regular journals. Despite this, most of the theoretical papers are worthwhile (I am not so sure about the experimental papers). The informal style of presentation makes them easy to assimilate and the reader can easily infer what the author really believes (unfortunately, the conventional style of journal writing often makes it hard to tell what parts of his work an author really takes seriously; perhaps we are too afraid that a frank style would give too many hostages to referees). This volume would therefore be a useful addition to libraries, although it is not the same essential work of reference as the proceedings of the one of the regular 'review conferences' (presumably it is beyond the reach of individuals at \$71.40).

C. H. Lewellyn Smith

C. H. Lewellyn Smith is Lecturer in Theoretical Physics at the University of Oxford, UK.

Physics and technology of lightning

Lightning. Vol. 1: Physics of Lightning. Pp. 496. Vol. 2: Lightning Protection. Pp. 353. Edited by R. H. Golde. (Academic: London, 1977.) Vol. 1 £21; vol. 2 £17.

In spite of the volume of observational data which has been acquired in the laboratory and the field, the process by which clouds become electrified is still not understood; neither is there any agreement on a physical model for the initiation and propagation of the lightning discharge. Two groups of workers have been principally involved in this research. The atmospheric physicist interests himself in the electrical processes and their relationship to the cloud environment, whereas the power engineer concentrates on the more specific task of lightning protection. Dr Golde has reflected this division by separating this collection of reviews on all aspects of lightning into two volumes entitled *The Physics of Lightning* and *Lightning Protection*.

The power engineer faces the practical problem of the protection of buildings, and power and telecommunications systems from the damaging effects of lightning. Given a knowledge of the time variation of the lightning current at the point of strike, he has developed a series of devices to minimise any damage caused. On an empirical level he has been very successful. Electricity supplies are only infrequently interrupted due to lightning, and the deaths due to lightning strikes during underground blasting have

been almost eliminated. Interference to telecommunications can be reduced to a tolerable level. Strikes to aircraft, and to microwave and television antennae will result in very little trouble being experienced if recommended practices are followed. The problem is mainly a matter of achieving maximum economic benefit. The chapters of the second volume comprise an impressive catalogue of the development and specifications of present techniques. Extensive reference to national and international recommendations should prove invaluable.

When we turn to the broader question of the physics of lightning itself, the situation is much less clear. Ideally for a complete description, we need to identify the microphysical and dynamical processes within clouds which lead to a separation of electric charge in agreement with the magnitude and position of the charges involved in cloud to ground and intra-cloud flashes. Next, we must explain the initiation and propagation of a conducting path within the cloud, leading finally to the observed current variation along the lightning channel and resultant electromagnetic radiation.

We are very far from such a comprehensive description. We have a vast array of observations from many parts of the world of parameters such as the time variation of lightning currents, leader velocities and the magnitude and heights of charges. It may well be that the lightning protection engineer with his demand for statistical information on frequency and intensity of discharges has resulted in many observers being pre-occupied by 'average' values of the above parameters which may vary over one or two orders of magnitude. Not surprisingly, it has proven impossible to fit the averages

together into a coherent description of the physical processes.

The editor has chosen acknowledged experts to deal with the different aspects, and each has produced an authoritative summary of the available information complete with exhaustive references. One omission is apparent. There is no chapter on the discharge process itself. Professor Allibone's chapter deals with the long spark between two electrodes, which is a practical problem in the insulation of electrical systems to withstand induced lightning transients. The lightning discharge is not initiated from a high capacity electrode but from within an initially insulating volume of cloud. In his 1969 book, Martin Uman described the physical models for leader and return stroke processes as being based on intuition rather than physics and being in a "lamentable" state. Should our lament continue? There is, for example, no discussion of the relative merits of the Schonland and Bruce models of the stepped leader in this book.

Controversy regarding the proposed charge separation mechanisms has been heated during recent years. It had for long been assumed that larger precipitation elements became negatively charged after interacting in some way with smaller cloud particles, and subsequent gravitational separation resulted in the observed high fields. This view has recently been challenged by the proposal that the convective motions within a cloud act like a giant Van der Graaf machine and that the precipitation development is incidental. It is refreshing to read the chapter by Professors Moore and Vonnegut carefully pointing out that the precipitation mechanism leads to contradictions and inconsistencies with observations, and suggesting that the convective mechanism could resolve some of these disagreements. There are few neutrals in this argument, so perhaps the best compromise would have been a second chapter giving the 'establishment' view of mechanisms, leaving the reader to judge.

During the past five years, we have seen the first steps towards a more ambitious form of lightning protection. Chaff has been dispensed within clouds in an effort to suppress lightning, and lightning has been triggered over land using rockets. It seems unlikely that there will be great progress in these endeavours until the basic problems of the physics of lightning are resolved. In their excellent chapter, Brook and Ogawa express the hope that a concerted effort using new techniques will lead to a far better understanding in a very few years. Meanwhile this first volume is a very useful summary of our present knowledge.

A. J. Illingworth

A. J. Illingworth is Lecturer in Physics at the University of Manchester Institute of Science and Technology, Manchester, UK.

Root system performance

Plant Root Systems: Their Function and Interaction with the Soil. By R. Scott Russell. Pp. 298. (McGraw-Hill: London, 1977.) £10.

THE past few years have seen the appearance of three large volumes dealing with plant roots, in particular with their structure, function and interaction with the soil. The appearance of these books at the same time is no coincidence: it reflects a growing awareness of the very important, though often indirect, effects root system performance can have on plant production. At the present time there is a place for a book which reviews much of the current knowledge represented (for example, by these large symposia volumes) and at the same time putting the results in a practical context. R. Scott-Russell's new book fulfils this role admirably.

It is appropriate that Scott-Russell should be the person to write such a book. As Director of the Letcombe Laboratory, he has guided much of the exciting and revealing work on roots and root systems which has been carried out in the past two decades. Many of the studies at the Letcombe Laboratory have attempted to link observations of roots and root system behaviour in manipulative laboratory experiments, with field observations; the ultimate aim of this work has been to highlight conditions which can limit root and subsequently crop growth.

Scott-Russell adopts a similar approach in his book, which has consequently been divided quite naturally into three parts. The first part describes root systems and their functions when no particular limit to root growth exists. This section, therefore, draws together much that is known about root system sizes in relation to above-ground parts, the growth and form that root systems take, the uptake of water and nutrients, and finally the relationships between roots and the rhizosphere flora.

The second part deals with the interaction between plant roots and soil. This section is introduced by a succinct chapter on the soil environment. Some might say too succinct but I feel that throughout the book the author has well realised where a fuller treatment would only serve to duplicate well known texts. Other chapters follow, covering some of the limiting conditions imposed in soils; chapters are therefore devoted to the effects of mechanical impedances on root growth and to the effects of anaerobic soils. The final chapter in this section deals with the soil/root interface and is

particularly useful as it brings together much of the work in this somewhat 'new' research area.

The final section of the book deals with soil tillage and particularly directs attention at the possibility and consequence of reducing the level of cultivation soils receive. This is a topical subject, as there is currently a considerable increase in areas of land receiving lower levels of cultivation (despite the contrary suggestion in Fig. 12:13 of the book). These chapters draw on earlier parts of the book and illustrate quite clearly how an understanding and use of agricultural practices can be assisted by a thorough knowledge of plant and soil processes, with which they interact.

I thoroughly recommend this book to anyone seeking to understand

current ideas on how plants function below ground. The text draws heavily from agricultural crop examples, perhaps because of the particular leanings of the author but more probably because there is much less data available on other plant types, for example, forest trees. Undergraduates in the plant and agricultural sciences would derive much benefit from this book, although its price might prove a deterrent. There are a few mistakes in the text but these are unimportant and unintrusive. The book is extremely well laid out with large clear print, and the use of different print types and print shades make reading very easy.

John M. Roberts

John M. Roberts is a Plant Physiologist working at the Institute of Hydrology, Wallingford, UK.

Lymphocyte structure and function

The Lymphocyte. Parts 1 and 2. Edited by J. J. Marchalonis. Pp. 1-369 and 373-704 (Marcel Dekker: Basel and New York, 1977.) SFr. 123 and 130.

IF definition is required, *The Lymphocyte* falls somewhere between a collection of reviews and a textbook; as such, is already obsolescent, as many of will perfectly fulfil neither role. As a collection of reviews, *The Lymphocyte* is already obsolescent, as many of the chapters were written four years ago. Judged as a textbook, it seems badly arranged. It has been chopped into somewhat arbitrary sections. Section three, for example, on "Physical Separation of Lymphocytes", contains but one short chapter of the same name. In the second volume, there is considerable overlap of material. Chapter 13, on "Ultrastructure of the Lymphocyte Surface", could quite readily have been omitted, as most of the material therein is given in other parts of the book. The text does, however, provide the student and those on the periphery of immunology with a wide range of information.

The ultrastructure of mammalian lymphocytes, and recirculating lymphocytes, are the two subjects of the first section of the book by T. E. Mandel and J. Sprent, respectively. Origins of lymphocytes and the lymphatic system are dealt with in the two chapters of the second section, one on the ontogenic emergence of immunocytes and the second on the phylogenetic emergence of lymphoid tissue and cells by N. Cohen. In view of his statement that "Crocodilians have big teeth and have been neglected by immunologists", Dr Cohen evidently feels that a dash of

humour in no way destroys good scientific literature.

The physical separation of lymphocytes by R. G. Miller is the short but nonetheless worthy topic of section three. The immune functions of lymphocytes are considered in section four which best achieves the object of the text since it gives a cross-section of current interests in the field. This section contains four chapters on such diverse topics as the influence of T cells on antibody formation and tolerance, cellular immune reactions, cell interactions in the *in vitro* immune response, and immune response genes.

The second volume of the book deals entirely with the lymphocyte plasma membrane, reflecting the major interest of the editor. The first chapter by Marchalonis himself deals with the purification of lymphocyte plasma membranes and the properties of lymphocyte surface immunoglobulins. This latter subject is also dealt with in two other chapters, one on lymphocyte surface antigens and the other on dynamic aspects of lymphocyte membranes. Other topics considered in this section include the lymphocyte coat or glycocalyx, lymphocyte membrane phospholipids, and mitogen stimulation of lymphocytes.

Thus, in summary, a wholehearted recommendation could not be given to this book partly because it is so dated and partly because better editing could have reduced its duplications. The publisher has mainly done a good job, although the quality of the figures is poor. The price, as so often happens these days, puts it into the grasp of the librarian and the reviewer rather than the private owner.

T. D. Heath

T. D. Heath is a research fellow in the Division of Biology of the Institute of Cancer Research, working at the Chester Beatty Research Institute, London, UK.

Ocean wave theories

The Dynamics of the Upper Ocean. (Second edition.) By O. M. Phillips. Pp. 336. (Cambridge University: Cambridge, London and New York, 1977.) £16.

JUST over a decade after its initial publication, Professor Phillips has produced a revised and up-to-date edition of his book. The book remains an applied mathematician's view of ocean surface waves, with an emphasis on the analytical mathematical methods that can be used. It also includes a chapter on internal waves and one on turbulence.

In the ten years before 1966, there had been some exciting developments in our understanding of sea waves. Stimulated initially by military requirements, methods of measuring wave spectra had been developed, and this led to studies of the statistical properties of sea waves and the evolution of the wave field with time. Storm waves from half way round the world could be identified, and useful predictive models of the wave field were being developed. There remained the problems of really understanding the physics of wave growth, the important non-linear interactions between waves, and the eventual loss of energy from waves by breaking and other mechanisms. But at that time, the solution of these problems could not have seemed too difficult.

The publication of the first edition was timely. It had the merit of covering a rapidly developing field in a uniform and scholarly manner. The mathematical content was also comprehensive and terse, making the book a suitable textbook. The main criticism was that in linking together the established ideas in the subject, the author occasionally slipped in some of his own ideas, which were suspect. However, despite this, the book has been, and is, a well merited success.

The second edition of the book has been extensively revised and a number of small sections added. The first third of the book is concerned with the equations of motion and the basic properties of surface waves. This is the best part of the book, being well written and containing elegant mathematics. New sections on wave action, wave-wave interactions and breaking waves, reflect the importance of these subjects.

With these preliminaries out of the way, chapter 4 (which takes up a third of the book) brings together Professor Phillips' ideas on how the sea wave spectra develop. Miles' and Phillips' theories of wave generation are developed at length, as in the first edition of the book, but Banner and Melville's recent observation of enhanced wind drag when a wave breaks is only mentioned briefly. This is surprising as the latter mechanism may well be more important than the traditional theories.

The section on wave-wave interactions is wrong, both in stating that the interactions are local in wave number space and also in seriously applying the analytical results valid for very narrow spectra to a realistic sea spectrum. As a consequence, the author cannot explain the observed enhancement of the peak of the spectrum and gets himself into a bit of a muddle.

Chapter 5 is concerned with internal waves. In the past decade, these have been studied in some detail, stimulated

by naval interest in the scintillation of sonar signals by internal waves. Garrett and Munk's internal wave spectra are included in the chapter, but the author is more successful at describing the mathematics of the problem than giving physical insight. The last chapter on turbulence is straightforward and contains more on experimental results than the first edition of the book.

Overall, I found the recent development of the subject reflected in this book, a little disappointing. Perhaps the theoreticians are a little too mesmerised by their own analytical methods and perhaps the experimenters were daunted by the cost of studying real sea waves. The subject may have its problems, but the elegant mathematics and coherent viewpoint still make this a useful book.

D. J. Webb

D. J. Webb is a Physical Oceanographer at the Institute of Oceanographic Sciences, Wormley, Godalming, UK.

Ferroelectricity and related topics

Principles and Applications of Ferroelectrics and Related Materials. By M. E. Lines and A. M. Glass. Pp. xiii+679. (Oxford University Press/Clarendon: London, New York and Oxford, 1977.) £29.

ONE of the interesting graphs in this book occurs in the preface and shows that the number of publications per year on ferroelectricity and related topics has increased exponentially from one in 1920 to one thousand in 1970—producing in the process such unlikely terms as "antiferrodistortive" for example, but the authors cannot be held responsible for that, and they successfully convinced me in fact that strontium titanate should not be called an antiferroelectric though ammonium dihydrogen phosphate should. They have been remarkably successful in covering almost every aspect of a subject which now involves most topics in solid-state physics, although they have required 680 pages to do so.

Brief introductory accounts are included of experimental techniques ranging from X-ray and neutron crystallography through nuclear and electron paramagnetic resonance to Mossbauer and optical spectroscopy, and of topics as formidable as the renormalisation group procedure in the theory of critical phenomena. Basic concepts are well explained without an excess of mathematical detail. Where the picture remains obscure,

as in the precise nature of soft modes in materials such as potassium dihydrogen phosphate or the explanation or explanations of the central peak in the energy spectrum of radiation scattered by materials near the critical temperature, there is no attempt to obscure the fact, and though I found a great deal that was new to me, I found very little with which I disagreed. My own prejudices are no doubt responsible for a feeling that the mean field theory is introduced too soon, with the exclusion of any discussion of the special role of coulomb interaction in the lattice dynamics of ionic crystals, but the book is on the whole so comprehensive and successful that it will be a rash author who sets out to write another one on this topic for many years to come.

It is perhaps surprising to find only 65 pages, less than 10% of the book, devoted to practical applications of ferroelectrics. This is not indicative of any lack of expertise on the part of the authors, both of whom are members of staff of Bell Telephone Laboratories, and Glass has published papers on pyroelectric detectors and electro-optic modulators, for example, but ferroelectrics have never found the range of application of semiconductors, although they are of comparable interest.

It is rare now to find a book of this scope with less than half a dozen authors and a couple of editors. Drs Lines and Glass have earned our congratulations on both scope and quality.

William Cochran

William Cochran is Professor of Natural Philosophy at the University of Edinburgh, Scotland.

● *Pulsars* by R. N. Manchester and J. H. Taylor (for review, see *Nature*, 19 January, 271, 285; 1978) has been published by Freeman (Reading) in the UK, price £14.50.